STIC Search Report
Biotech-Chem Library

STIC Database Tracking Number: 116491

TO: Jeffrey Parkin
Location: REM-3D39/3C18
Art Unit: 1648
Monday, March 15, 2004
Case Serial Number: 10/055524

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: (571)272-2527

paul.schulwitz@uspto.gov

Search Notes

Examiner Parkin,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Paul Schulwitz
Technical Information Specialist
STIC Biotech/Chem Library
(571)272-2527

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 12, 2004, 14:22:45 ; Search time 54 Seconds
(without alignments)
52.324 Million cell updates/sec

Title: PARKIN524.PEP
Perfect score: 46
Sequence: 1 kpvvstqlll 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	10	4	ABP16652	Abp16652 HIV B07 s
2	46	100.0	10	6	ABU69893	Abu69893 Human imm
3	46	100.0	10	6	ABU70212	Abu70212 Human imm
4	46	100.0	10	6	ABU70010	Abu70010 Human imm
5	46	100.0	10	7	ADD96745	Add96745 HIV-1 cro
6	46	100.0	10	7	ADD96543	Add96543 HIV-1 cro
7	46	100.0	10	7	ADD96426	Add96426 HIV-1 cro
8	46	100.0	11	4	ABP17099	Abp17099 HIV B27 s
9	46	100.0	15	4	ABP24380	Abp24380 HIV DR su
10	46	100.0	15	4	ABP24379	Abp24379 HIV DR su
11	46	100.0	15	4	ABP24416	Abp24416 HIV DR su
12	46	100.0	229	5	AAM48947	Aam48947 HIV-1 sub
13	46	100.0	236	5	AAE15742	Aae15742 Plasmid C
14	46	100.0	417	2	AAW43071	Aaw43071 HIV-1 gp1
15	46	100.0	474	2	AAW37067	Aaw37067 HIV-1 bre
16	46	100.0	474	2	AAW37066	Aaw37066 HIV-1 bre
17	46	100.0	478	7	AAE39557	Aae39557 HIV gp120
18	46	100.0	478	7	AAE39558	Aae39558 HIV gp120
19	46	100.0	498	2	AAW37055	Aaw37055 HIV-1 bre
20	46	100.0	498	2	AAW37054	Aaw37054 HIV-1 bre
21	46	100.0	501	7	AAE39544	Aae39544 HIV gp120
22	46	100.0	501	7	AAE39543	Aae39543 HIV gp120
23	46	100.0	586	2	AAR49855	Aar49855 Sequence
24	46	100.0	587	2	AAY29900	Aay29900 Human MDC
25	46	100.0	591	2	AAR49856	Aar49856 Sequence
26	46	100.0	595	2	AAY29902	Aay29902 Human MCP
27	46	100.0	601	2	AAY29901	Aay29901 Human IP-
28	46	100.0	651	5	AAU84397	Aau84397 HIV ENV c
29	46	100.0	842	3	AAB69350	Aab69350 HIV-1 non
30	46	100.0	845	3	AAB69349	Aab69349 HIV-1 non
31	46	100.0	846	3	AAB69353	Aab69353 HIV-1 non
32	46	100.0	846	3	AAB69345	Aab69345 HIV-1 non
33	46	100.0	849	3	AAB69346	Aab69346 HIV-1 non
34	46	100.0	849	3	AAY96945	Aay96945 HIV synth
35	46	100.0	853	4	AAB82762	Aab82762 Ancestral
36	46	100.0	854	5	ABB06214	Abb06214 HIV Env i
37	46	100.0	854	6	ABU66568	Abu66568 Human imm
38	46	100.0	854	6	ABR55687	Abr55687 HIV isola
39	46	100.0	854	7	ADC13221	Adc13221 Protein o
40	46	100.0	855	3	AAB69351	Aab69351 HIV-1 non
41	46	100.0	856	7	ABB80283	Abb80283 gp41 of 1
42	46	100.0	857	2	AAR12261	Aar12261 HIV-1 str
43	46	100.0	857	3	AAB69355	Aab69355 HIV-1 non
44	46	100.0	858	5	AAM48951	Aam48951 HIV-1 sub
45	46	100.0	859	1	AAP81865	Aap81865 Sequence

ALIGNMENTS

RESULT 1
ABP16652
ID ABP16652 standard; peptide; 10 AA.
XX
AC ABP16652;
XX
DT 11-SEP-2003 (revised)
DT 15-JUL-2002 (first entry)
XX
DE HIV B07 super motif env peptide #32.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
KW vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US027766.
XX
PR 05-OCT-1999; 99US-00412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1.
XX
PS Claim 32; Page 209; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
CC be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or

```
CC   tumour-associated antigens, which therefore reduces the likelihood of
CC   escape mutants. Furthermore, immunosuppressive groups that may be present
CC   in whole antigens can be avoided with the use of group-based vaccines. An
CC   additional advantage of an group-based vaccine approach is the ability to
CC   combine selected groups (CTL and HTL), and further, to modify the
CC   composition of the groups, achieving, for example, enhanced
CC   immunogenicity. Accordingly, the immune response can be modulated, as
CC   appropriate, for the target disease. Similar engineering of the response
CC   is not possible with traditional approaches. ABP11501 to ABP25412
CC   represent peptide sequences used in the exemplification of the present
CC   invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 46;  DB 4;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 2
ABU69893
ID   ABU69893 standard; peptide; 10 AA.
XX
AC   ABU69893;
XX
DT   23-OCT-2003  (revised)
DT   05-JUN-2003  (first entry)
XX
DE   Human immunodeficiency virus 1 (HIV1) vaccine candidate peptide #211.
XX
KW   Human immunodeficiency virus; HIV; vaccine; immunological excipient;
KW   anti-HIV immune response; T cell response;
KW   viral multiplication inhibitor; chronic viraemia; AIDS.
XX
OS   Human immunodeficiency virus 1.
XX
PN   US2002182222-A1.
XX
PD   05-DEC-2002.
XX
PF   26-OCT-2001; 2001US-00055524.
XX
PR   10-JUL-1998;   98US-0092346P.
PR   08-JAN-1999;   99US-0115145P.
PR   23-APR-1999;   99US-0130677P.
PR   09-JUL-1999;   99US-00351036.
XX
PA   (GROO/) GROOT A D.
XX
PI   Groot AD;
XX
DR   WPI; 2003-352642/33.
XX
PT   New vaccine comprising human immunodeficiency virus (HIV) vaccine
PT   candidate peptides, useful as antigens for raising anti-HIV immune
PT   responses, such as T cell responses, and for inducing antibodies and
PT   impairing viral multiplication.
XX
PS   Claim 1; Page 15; 32pp; English.
XX
CC   The invention describes a vaccine comprising a human immunodeficiency
CC   virus (HIV) vaccine candidate peptide containing an amino acid sequence
CC   selected from 669 amino acid sequences given in the specification, in an
CC   immunological excipient. The HIV vaccine peptides are useful as antigens
CC   for raising anti-HIV immune responses, such as T cell responses, and for
CC   inducing antibodies that react with HIV-1 and impairing viral
CC   multiplication in vivo. These antibodies reduce viral multiplication
CC   during any initial acute infection with HIV-1 and minimise chronic
CC   viraemia or progression leading to AIDS. This is the amino acid sequence
CC   of a human immunodeficiency virus 1 (HIV1) vaccine candiate peptide.
CC   (Updated on 23-OCT-2003 to standardise OS field)
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 46;  DB 6;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 3
ABU70212
ID   ABU70212 standard; peptide; 10 AA.
XX
AC   ABU70212;
XX
DT   23-OCT-2003  (revised)
DT   05-JUN-2003  (first entry)
XX
DE   Human immunodeficiency virus 1 (HIV1) vaccine candidate peptide #530.
XX
KW   Human immunodeficiency virus; HIV; vaccine; immunological excipient;
KW   anti-HIV immune response; T cell response;
KW   viral multiplication inhibitor; chronic viraemia; AIDS.
XX
OS   Human immunodeficiency virus 1.
XX
PN   US2002182222-A1.
XX
PD   05-DEC-2002.
XX
PF   26-OCT-2001; 2001US-00055524.
XX
PR   10-JUL-1998;   98US-0092346P.
PR   08-JAN-1999;   99US-0115145P.
PR   23-APR-1999;   99US-0130677P.
PR   09-JUL-1999;   99US-00351036.
XX
PA   (GROO/) GROOT A D.
XX
PI   Groot AD;
XX
DR   WPI; 2003-352642/33.
XX
PT   New vaccine comprising human immunodeficiency virus (HIV) vaccine
PT   candidate peptides, useful as antigens for raising anti-HIV immune
PT   responses, such as T cell responses, and for inducing antibodies and
PT   impairing viral multiplication.
XX
PS   Claim 1; Page 19; 32pp; English.
XX
CC   The invention describes a vaccine comprising a human immunodeficiency
CC   virus (HIV) vaccine candidate peptide containing an amino acid sequence
CC   selected from 669 amino acid sequences given in the specification, in an
CC   immunological excipient. The HIV vaccine peptides are useful as antigens
CC   for raising anti-HIV immune responses, such as T cell responses, and for
CC   inducing antibodies that react with HIV-1 and impairing viral
CC   multiplication in vivo. These antibodies reduce viral multiplication
CC   during any initial acute infection with HIV-1 and minimise chronic
CC   viraemia or progression leading to AIDS. This is the amino acid sequence
CC   of a human immunodeficiency virus 1 (HIV1) vaccine candiate peptide.
CC   (Updated on 23-OCT-2003 to standardise OS field)
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 46;  DB 6;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 4
ABU70010
ID   ABU70010 standard; peptide; 10 AA.
XX
AC   ABU70010;
XX
DT   23-OCT-2003  (revised)
DT   05-JUN-2003  (first entry)
XX
DE   Human immunodeficiency virus 1 (HIV1) vaccine candidate peptide #328.
XX
KW   Human immunodeficiency virus; HIV; vaccine; immunological excipient;
KW   anti-HIV immune response; T cell response;
KW   viral multiplication inhibitor; chronic viraemia; AIDS.
XX
OS   Human immunodeficiency virus 1.
XX
PN   US2002182222-A1.
XX
PD   05-DEC-2002.
XX
PF   26-OCT-2001; 2001US-00055524.
XX
PR   10-JUL-1998;   98US-0092346P.
PR   08-JAN-1999;   99US-0115145P.
PR   23-APR-1999;   99US-0130677P.
PR   09-JUL-1999;   99US-00351036.
XX
PA   (GROO/) GROOT A D.
XX
PI   Groot AD;
XX
DR   WPI; 2003-352642/33.
XX
PT   New vaccine comprising human immunodeficiency virus (HIV) vaccine
PT   candidate peptides, useful as antigens for raising anti-HIV immune
PT   responses, such as T cell responses, and for inducing antibodies and
PT   impairing viral multiplication.
XX
PS   Claim 1; Page 16; 32pp; English.
XX
CC   The invention describes a vaccine comprising a human immunodeficiency
CC   virus (HIV) vaccine candidate peptide containing an amino acid sequence
CC   selected from 669 amino acid sequences given in the specification, in an
CC   immunological excipient. The HIV vaccine peptides are useful as antigens
CC   for raising anti-HIV immune responses, such as T cell responses, and for
CC   inducing antibodies that react with HIV-1 and impairing viral
CC   multiplication in vivo. These antibodies reduce viral multiplication
CC   during any initial acute infection with HIV-1 and minimise chronic
CC   viraemia or progression leading to AIDS. This is the amino acid sequence
CC   of a human immunodeficiency virus 1 (HIV1) vaccine candiate peptide.
CC   (Updated on 23-OCT-2003 to standardise OS field)
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 46;  DB 6;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 5
ADD96745
ID   ADD96745 standard; peptide; 10 AA.
XX
AC   ADD96745;
XX
DT   29-JAN-2004  (first entry)
XX
DE   HIV-1 cross-clade candidate peptide #530.
XX
KW   HIV-1; cross-clade candidate peptide; HIV clade;
KW   major histocompatibility complex; MHC; human leukocyte antigen; HLA;
KW   T-cell activation; HIV positive patient; HIV infection; anti-HIV.
XX
OS   Human immunodeficiency virus 1.
XX
PN   US2003180314-A1.
XX
PD   25-SEP-2003.
XX
PF   22-JUL-2002; 2002US-00200708.
XX
PR   10-JUL-1998;   98US-0092346P.
PR   08-JAN-1999;   99US-0115145P.
PR   23-APR-1999;   99US-0130677P.
PR   09-JUL-1999;   99US-00351036.
XX
PA   (DEGR/) DEGROOT A.
XX
PI   Degroot A;
XX
DR   WPI; 2003-852210/79.
XX
PT   New cross-clade HIV candidate peptide that binds a human major
PT   histocompatibility complex binding matrix motif or activates T-cells from
PT   HIV positive patients, useful for preventing or treating HIV infection.
XX
PS   Example 3; SEQ ID NO 530; 146pp; English.
XX
CC   The present invention relates to HIV-1 cross-clade candidate peptides
CC   comprising a sequence of about 8-50 amino acids, the sequence having
CC   complete, sequential sequence identity with a partial HIV-1 amino acid
CC   sequence that is absolutely conserved across at least 2 clades of HIV.
CC   The HIV-1 cross-clade candidate peptides possess at least one of the
CC   biological properties selected from (i) the ability to bind a human major
CC   histocompatibility complex (MHC) binding matrix motif for a human MHC
CC   allele, (ii) the ability to bind MHC human leukocyte antigen (HLA) in the
CC   T2 in vitro peptide binding assay, and (iii) the ability to activate T-
CC   cells from HIV positive patients in at least one in vitro assay. The
CC   invention also discloses a pharmaceutical composition comprising the
CC   above peptide, and methods for the production and use of the cross-clade
CC   peptides. The composition and methods are useful in preventing or
CC   treating HIV infection. The present sequence represents a HIV-1 cross-
CC   clade candidate peptide.
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 46;  DB 7;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 6
ADD96543
ID   ADD96543 standard; peptide; 10 AA.
XX
AC   ADD96543;
XX
DT   29-JAN-2004  (first entry)
XX
DE   HIV-1 cross-clade candidate peptide #328.
XX
KW   HIV-1; cross-clade candidate peptide; HIV clade;
```

```
KW   major histocompatibility complex; MHC; human leukocyte antigen; HLA;
KW   T-cell activation; HIV positive patient; HIV infection; anti-HIV.
XX
OS   Human immunodeficiency virus 1.
XX
PN   US2003180314-A1.
XX
PD   25-SEP-2003.
XX
PF   22-JUL-2002; 2002US-00200708.
XX
PR   10-JUL-1998;   98US-0092346P.
PR   08-JAN-1999;   99US-0115145P.
PR   23-APR-1999;   99US-0130677P.
PR   09-JUL-1999;   99US-00351036.
XX
PA   (DEGR/) DEGROOT A.
XX
PI   Degroot A;
XX
DR   WPI; 2003-852210/79.
XX
PT   New cross-clade HIV candidate peptide that binds a human major
PT   histocompatibility complex binding matrix motif or activates T-cells from
PT   HIV positive patients, useful for preventing or treating HIV infection.
XX
PS   Example 3; SEQ ID NO 328; 146pp; English.
XX
CC   The present invention relates to HIV-1 cross-clade candidate peptides
CC   comprising a sequence of about 8-50 amino acids, the sequence having
CC   complete, sequential sequence identity with a partial HIV-1 amino acid
CC   sequence that is absolutely conserved across at least 2 clades of HIV.
CC   The HIV-1 cross-clade candidate peptides possess at least one of the
CC   biological properties selected from (i) the ability to bind a human major
CC   histocompatibility complex (MHC) binding matrix motif for a human MHC
CC   allele, (ii) the ability to bind MHC human leukocyte antigen (HLA) in the
CC   T2 in vitro peptide binding assay, and (iii) the ability to activate T-
CC   cells from HIV positive patients in at least one in vitro assay. The
CC   invention also discloses a pharmaceutical composition comprising the
CC   above peptide, and methods for the production and use of the cross-clade
CC   peptides. The composition and methods are useful in preventing or
CC   treating HIV infection. The present sequence represents a HIV-1 cross-
CC   clade candidate peptide.
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 46;  DB 7;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 7
ADD96426
ID   ADD96426 standard; peptide; 10 AA.
XX
AC   ADD96426;
XX
DT   29-JAN-2004  (first entry)
XX
DE   HIV-1 cross-clade candidate peptide #211.
XX
KW   HIV-1; cross-clade candidate peptide; HIV clade;
KW   major histocompatibility complex; MHC; human leukocyte antigen; HLA;
KW   T-cell activation; HIV positive patient; HIV infection; anti-HIV.
XX
OS   Human immunodeficiency virus 1.
XX
PN   US2003180314-A1.
XX
PD   25-SEP-2003.
XX
PF   22-JUL-2002; 2002US-00200708.
XX
PR   10-JUL-1998;   98US-0092346P.
PR   08-JAN-1999;   99US-0115145P.
PR   23-APR-1999;   99US-0130677P.
PR   09-JUL-1999;   99US-00351036.
XX
PA   (DEGR/) DEGROOT A.
XX
PI   Degroot A;
XX
DR   WPI; 2003-852210/79.
XX
PT   New cross-clade HIV candidate peptide that binds a human major
PT   histocompatibility complex binding matrix motif or activates T-cells from
PT   HIV positive patients, useful for preventing or treating HIV infection.
XX
PS   Example 3; SEQ ID NO 211; 146pp; English.
XX
CC   The present invention relates to HIV-1 cross-clade candidate peptides
CC   comprising a sequence of about 8-50 amino acids, the sequence having
CC   complete, sequential sequence identity with a partial HIV-1 amino acid
CC   sequence that is absolutely conserved across at least 2 clades of HIV.
CC   The HIV-1 cross-clade candidate peptides possess at least one of the
CC   biological properties selected from (i) the ability to bind a human major
CC   histocompatibility complex (MHC) binding matrix motif for a human MHC
CC   allele, (ii) the ability to bind MHC human leukocyte antigen (HLA) in the
CC   T2 in vitro peptide binding assay, and (iii) the ability to activate T-
CC   cells from HIV positive patients in at least one in vitro assay. The
CC   invention also discloses a pharmaceutical composition comprising the
CC   above peptide, and methods for the production and use of the cross-clade
CC   peptides. The composition and methods are useful in preventing or
CC   treating HIV infection. The present sequence represents a HIV-1 cross-
CC   clade candidate peptide.
XX
SQ   Sequence 10 AA;

  Query Match             100.0%;  Score 46;  DB 7;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 8
ABP17099
ID   ABP17099 standard; peptide; 11 AA.
XX
AC   ABP17099;
XX
DT   11-SEP-2003  (revised)
DT   15-JUL-2002  (first entry)
XX
DE   HIV B27 super motif env peptide #124.
XX
KW   HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
KW   vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
KW   vaccine; HIV infection; immunisation; virucide.
XX
OS   Human immunodeficiency virus 1.
XX
PN   WO200124810-A1.
XX
PD   12-APR-2001.
XX
PF   05-OCT-2000; 2000WO-US027766.
XX
PR   05-OCT-1999;   99US-00412863.
XX
```

PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1.
XX
PS Claim 32; Page 219; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
CC be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines. An
CC additional advantage of an group-based vaccine approach is the ability to
CC combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 46; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 2 KPVVSTQLLL 11

RESULT 9
ABP24380
ID ABP24380 standard; peptide; 15 AA.
XX
AC ABP24380;
XX
DT 11-SEP-2003 (revised)
DT 15-JUL-2002 (first entry)
XX
DE HIV DR super motif env peptide #2.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
KW vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US027766.
XX
PR 05-OCT-1999; 99US-00412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1.
XX
PS Claim 32; Page 369; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
CC be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines. An
CC additional advantage of an group-based vaccine approach is the ability to
CC combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 46; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 2 KPVVSTQLLL 11

RESULT 10
ABP24379
ID ABP24379 standard; peptide; 15 AA.
XX
AC ABP24379;
XX
DT 11-SEP-2003 (revised)
DT 15-JUL-2002 (first entry)
XX
DE HIV DR super motif env peptide #1.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
KW vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US027766.
XX
PR 05-OCT-1999; 99US-00412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;

XX
DR WPI; 2001-354887/37.
XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1.
XX
PS Claim 32; Page 369; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
CC be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines. An
CC additional advantage of an group-based vaccine approach is the ability to
CC combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 46; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 1 KPVVSTQLLL 10

RESULT 11
ABP24416
ID ABP24416 standard; peptide; 15 AA.
XX
AC ABP24416;
XX
DT 11-SEP-2003 (revised)
DT 15-JUL-2002 (first entry)
XX
DE HIV DR super motif env peptide #38.
XX
KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;
KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;
KW vaccine; HIV infection; immunisation; virucide.
XX
OS Human immunodeficiency virus 1.
XX
PN WO200124810-A1.
XX
PD 12-APR-2001.
XX
PF 05-OCT-2000; 2000WO-US027766.
XX
PR 05-OCT-1999; 99US-00412863.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-354887/37.

XX
PT Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)
PT peptide groups, useful for vaccinating against HIV-1.
XX
PS Claim 32; Page 369; 448pp; English.
XX
CC The present invention describes a composition (I) comprising a prepared
CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid
CC sequence selected from 51 defined amino acid sequences (ABL25347 to
CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may
CC be used for immunising subjects against HIV-1 infections. The use of
CC group-based vaccines has several advantages over traditional vaccines,
CC particularly when compared to the use of whole antigens in vaccine
CC compositions. There is evidence that the immune response to whole
CC antigens is directed largely toward variable regions of the antigen,
CC allowing for immune escape due to mutations. The groups for inclusion in
CC an group-based vaccine may be selected from conserved regions of viral or
CC tumour-associated antigens, which therefore reduces the likelihood of
CC escape mutants. Furthermore, immunosuppressive groups that may be present
CC in whole antigens can be avoided with the use of group-based vaccines. An
CC additional advantage of an group-based vaccine approach is the ability to
CC combine selected groups (CTL and HTL), and further, to modify the
CC composition of the groups, achieving, for example, enhanced
CC immunogenicity. Accordingly, the immune response can be modulated, as
CC appropriate, for the target disease. Similar engineering of the response
CC is not possible with traditional approaches. ABP11501 to ABP25412
CC represent peptide sequences used in the exemplification of the present
CC invention. (Updated on 11-SEP-2003 to standardise OS field)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 46; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 5 KPVVSTQLLL 14

RESULT 12
AAM48947
ID AAM48947 standard; protein; 229 AA.
XX
AC AAM48947;
XX
DT 29-AUG-2003 (revised)
DT 19-APR-2002 (first entry)
XX
DE HIV-1 subtype C env protein fragment consensus sequence.
XX
KW HIV-1 subtype C; vaccine; HIV infection; AIDS; env; antiviral.
XX
OS Human immunodeficiency virus; type I.
XX
PN WO200204494-A2.
XX
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-IB001208.
XX
PR 07-JUL-2000; 2000US-0216995P.
PR 10-JUL-2000; 2000ZA-00003437.
PR 15-SEP-2000; 2000ZA-00004924.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
PA (UYCA-) UNIV CAPE TOWN.
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Williamson C, Swanstrom RI, Morris L, Karim SA, Johnston RE;
XX
DR WPI; 2002-171700/22.
XX

```
PT   Selecting HIV-1 subtype C isolates, which are useful in developing
PT   vaccines against HIV infection, comprises isolating viruses with high
PT   sequence identity to a consensus sequence whose phenotype is associated
PT   with the HIV subtype.
XX
PS   Claim 26; Page 37; 69pp; English.
XX
CC   The present invention relates to a process for the selection of human
CC   immunodeficiency virus (HIV) subtype isolates for use in the development
CC   of a prophylactic and/or therapeutic pharmaceutical composition. The
CC   process involves selecting isolated virus or viruses with a high sequence
CC   identity to a consensus sequence and a phenotype which is associated with
CC   transmission for the particular HIV subtype. The composition can be used
CC   in the production of vaccines against HIV. The present sequence is a
CC   consensus sequence for the HIV-1 subtype C env protein. (Updated on 29-
CC   AUG-2003 to standardise OS field)
XX
SQ   Sequence 229 AA;

  Query Match             100.0%;  Score 46;  DB 5;  Length 229;
  Best Local Similarity   100.0%;  Pred. No. 0.29;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db         36 KPVVSTQLLL 45


RESULT 13
AAE15742
ID   AAE15742 standard; protein; 236 AA.
XX
AC   AAE15742;
XX
DT   29-AUG-2003  (revised)
DT   26-MAR-2002  (first entry)
XX
DE   Plasmid CMV-HIV UB-#23 DNA encoded protein.
XX
KW   HIV; human immunodeficiency virus; CMV; cytomegalovirus; cytostatic;
KW   immunosuppressive; virucide; antibacterial; fungicide; protozoacide;
KW   antirheumatic; antiinflammatory; antiarthritic; neuroprotective;
KW   rheumatoid arthritis; cancer; multiple sclerosis; immune response;
KW   vasotropic; vaccine; gene therapy; autoimmune disease; vasculitis;
KW   ubiquitin.
XX
OS   Human immunodeficiency virus.
OS   Cytomegalovirus.
OS   Chimeric.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 46
FT                   /note= "Encoded by AGG"
FT   Misc-difference 69
FT                   /note= "Encoded by GTC"
FT   Misc-difference 71
FT                   /note= "Encoded by TTC"
FT   Misc-difference 72
FT                   /note= "Encoded by ATG"
FT   Misc-difference 75
FT                   /note= "Encoded by GCT"
FT   Misc-difference 84
FT                   /note= "Encoded by ACA"
FT   Misc-difference 94
FT                   /note= "Encoded by AGC"
FT   Misc-difference 98
FT                   /note= "Encoded by ATA"
FT   Misc-difference 110
FT                   /note= "Encoded by GCA"
FT   Misc-difference 114
FT                   /note= "Encoded by ATG"
FT   Misc-difference 119
FT                   /note= "Encoded by GAT"
FT   Misc-difference 147
FT                   /note= "Encoded by TTT"
FT   Misc-difference 163
FT                   /note= "Encoded by CCA"
FT   Misc-difference 223
FT                   /note= "Encoded by AAA"
XX
PN   WO200176643-A1.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-US011372.
XX
PR   07-APR-2000; 2000US-0195680P.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Orson FM,  Kinsey BM,  Bhogal BS;
XX
DR   WPI; 2002-066308/09.
DR   N-PSDB; AAD25513.
XX
PT   Composition for oral delivery of vaccines, comprises expression vector
PT   containing antigenic genomic sequence, bound to aggregated protein-
PT   polycationic polymer conjugate or suspension.
XX
PS   Example 10; Page 75-76; 145pp; English.
XX
CC   The invention relates to a composition comprising an expression vector
CC   bound to an aggregated protein-polycationic polymer conjugate or
CC   suspension. The expression vector contains a promoter polynucleotide
CC   sequence operatively linked to a polynucleotide sequence encoding an
CC   antigen which is a fragment of a gene or genome associated with an
CC   infectious disease, cancer and autoimmune disease such as rheumatoid
CC   arthritis, vasculitis, and multiple sclerosis, pathogenic genomes
CC   consisting of bacterium, fungus, protozoa and virus such as human
CC   immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C
CC   virus (HCV), influenza and respiratory syncytial virus (RSV), and
CC   optionally comprising a nucleotide sequence encoding a cytokine (or a
CC   cytokine expression vector), is useful for inducing an immune response
CC   (systemic and/or mucosal) in an organism. The cytokine expression vector
CC   contains a sequence for granulocyte macrophage-colony stimulating factor
CC   (GM-CSF) or interleukin-12 (IL-12). The polynucleotide sequences encoding
CC   the antigen and the cytokine are under transcriptional control of same or
CC   different promoter polynucleotide sequences. The expression vector, as a
CC   DNA vaccine is useful for treating a condition in an organism. The
CC   present sequence is protein encoded by plasmid CMV-HIV UB-#23 DNA. The
CC   protein contains ubiquitin fused to a protein fragment containing the
CC   immunodominant epitope for gp120, related to the invention. (Updated on
CC   29-AUG-2003 to standardise OS field)
XX
SQ   Sequence 236 AA;

  Query Match             100.0%;  Score 46;  DB 5;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db         46 KPVVSTQLLL 55


RESULT 14
AAW43071
ID   AAW43071 standard; peptide; 417 AA.
XX
AC   AAW43071;
XX
DT   17-OCT-2003  (revised)
DT   11-SEP-1998  (first entry)
XX
DE   HIV-1 gp120 protein fragment from isolate NY5.
XX
```

```
KW   gp120 protein; purification; fractionation; ion exchange; chromatography;
KW   binding affinity; CD4; hydrophobic interaction; size exclusion; vaccine.
XX
OS   Human immunodeficiency virus 1.
XX
PN   US5696238-A.
XX
PD   09-DEC-1997.
XX
PF   11-MAY-1995;   95US-00439286.
XX
PR   20-AUG-1991;   91US-00684963.
PR   16-AUG-1993;   93US-00109002.
PR   09-MAY-1994;   94US-00240073.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Scandella C,  Haigwood NL;
XX
DR   WPI; 1998-041353/04.
XX
PT   Purification of HIV gp120 - using chromatographic methods.
XX
PS   Disclosure; Fig 2A-W; 53pp; English.
XX
CC   AAW43066-W43080 are fragments of the gp120 protein from different human
CC   immunodeficiency virus type I (HIV-1) isolates. These proteins are used
CC   in a novel method for purifying HIV gp120 so as to provide a purified
CC   gp120 glycopeptide having protein/protein binding properties
CC   substantially identical to natural viral HIV gp120. The method involves
CC   fractionating a crude gp120 preparation containing full-length,
CC   glycosylated gp120 using ion exchange chromatography so as to provide a
CC   first collection of fractions. A a fraction from the first collection is
CC   selected that exhibits specific binding affinity for CD4 peptide, thereby
CC   producing a first fractionated material. The first fractionated material
CC   is fractionated by hydrophobic interaction chromatography so as to
CC   provide a second collection of fractions from which a second collection
CC   is selected that exhibits specific binding affinity for CD4 peptide. This
CC   second fraction is fractionated by size exclusion chromatography so as to
CC   provide a third collection of fractions exhibiting specific binding
CC   affinity for CD4 peptide, thereby providing the purified gp120. The
CC   purified gp120 can be used for antibody production and in vaccines.
CC   (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ   Sequence 417 AA;

  Query Match             100.0%;  Score 46;  DB 2;  Length 417;
  Best Local Similarity   100.0%;  Pred. No. 0.56;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        246 KPVVSTQLLL 255


RESULT 15
AAW37067
ID   AAW37067 standard; protein; 474 AA.
XX
AC   AAW37067;
XX
DT   17-OCT-2003  (revised)
DT   20-JUL-1998  (first entry)
XX
DE   HIV-1 breakthrough isolate clone C17.3 gp120 polypeptide.
XX
KW   HIV-1; envelope protein; gp120; MN-rgp120; vaccine; AIDS.
XX
OS   Human immunodeficiency virus 1.
XX
PN   WO9801564-A1.
XX
PD   15-JAN-1998.
XX
PF   03-JUL-1997;   97WO-US009690.
XX
PR   08-JUL-1996;   96US-0676737P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Berman PW;
XX
DR   WPI; 1998-101059/09.
DR   N-PSDB; AAV00530.
XX
PT   Human immunodeficiency virus gp120 sequences from vaccine breakthrough
PT   strains - useful in providing added protection against HIV not provided
PT   by original vaccine.
XX
PS   Claim 1; Page 145-148; 193pp; English.
XX
CC   This is the deduced sequence of a gp120 envelope polypeptide of clone
CC   C17.3 of a HIV-1 breakthrough isolate obtained from an individual
CC   infected with HIV-1 through high risk activity while participating in
CC   Phase I or Phase 2 trials of MN-rgp120, a candidate recombinant gp120 HIV
CC   -1 vaccine. Nucleotide sequences (see AAV00517-30) for gp120 polypeptides
CC   (see AAW37054-67) were obtained from 2 clones of each of 7 breakthrough
CC   isolates. All 7 envelope glycoproteins have sequences typical of subtype
CC   (clade) B viruses. The overall homology with MN-rgp120 is 69-80%. Use of
CC   the gp120 polypeptides from one or more of the breakthrough isolates,
CC   usually together with MN-rgp120, can provide protection against HIV
CC   strains that are sufficiently different from the vaccine strain (e.g. MN-
CC   rgp120) that the vaccine does not confer protection against those
CC   strains. The gp120 proteins can also be used in screening assays to
CC   identify antagonists of CC-CKR chemokine receptors. Antibodies induced by
CC   the polypeptides are also provided. (Updated on 17-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 474 AA;

  Query Match             100.0%;  Score 46;  DB 2;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 0.65;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        207 KPVVSTQLLL 216


Search completed: March 12, 2004, 14:23:56
Job time : 57 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 12, 2004, 14:22:50 ; Search time 23 Seconds
(without alignments)
22.446 Million cell updates/sec

Title: PARKIN524.PEP
Perfect score: 46
Sequence: 1 kpvvstqlll 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	474	3	US-08-889-841B-36	Sequence 36, Appl
2	46	100.0	474	3	US-08-889-841B-39	Sequence 39, Appl
3	46	100.0	474	4	US-09-419-362-36	Sequence 36, Appl
4	46	100.0	474	4	US-09-419-362-39	Sequence 39, Appl
5	46	100.0	498	3	US-08-889-841B-2	Sequence 2, Appli
6	46	100.0	498	3	US-08-889-841B-5	Sequence 5, Appli
7	46	100.0	498	4	US-09-419-362-2	Sequence 2, Appli
8	46	100.0	498	4	US-09-419-362-5	Sequence 5, Appli
9	46	100.0	587	4	US-09-646-028-50	Sequence 50, Appl
10	46	100.0	595	4	US-09-646-028-56	Sequence 56, Appl
11	46	100.0	596	2	US-08-392-806A-2	Sequence 2, Appli
12	46	100.0	596	3	US-09-257-490-2	Sequence 2, Appli
13	46	100.0	600	2	US-08-392-806A-4	Sequence 4, Appli
14	46	100.0	600	3	US-09-257-490-4	Sequence 4, Appli
15	46	100.0	601	4	US-09-646-028-52	Sequence 52, Appl
16	46	100.0	613	3	US-09-257-490-14	Sequence 14, Appl
17	46	100.0	855	4	US-09-206-551-11	Sequence 11, Appl
18	46	100.0	855	4	US-09-206-551-12	Sequence 12, Appl
19	43	93.5	15	4	US-09-009-953-269	Sequence 269, App
20	43	93.5	20	1	US-08-218-025A-52	Sequence 52, Appl
21	43	93.5	21	1	US-08-218-025A-147	Sequence 147, App
22	43	93.5	21	3	US-09-220-389A-1	Sequence 1, Appli
23	43	93.5	22	1	US-08-218-025A-110	Sequence 110, App
24	43	93.5	22	1	US-08-218-025A-143	Sequence 143, App
25	43	93.5	22	2	US-08-493-235-19	Sequence 19, Appl
26	43	93.5	84	4	US-09-536-977-10	Sequence 10, Appl
27	43	93.5	425	4	US-09-536-977-46	Sequence 46, Appl
28	43	93.5	425	4	US-09-536-977-48	Sequence 48, Appl
29	43	93.5	425	4	US-09-536-977-50	Sequence 50, Appl
30	43	93.5	425	4	US-09-536-977-52	Sequence 52, Appl
31	43	93.5	455	3	US-08-889-841B-46	Sequence 46, Appl
32	43	93.5	455	4	US-09-419-362-46	Sequence 46, Appl
33	43	93.5	469	3	US-08-889-841B-23	Sequence 23, Appl
34	43	93.5	469	4	US-09-419-362-23	Sequence 23, Appl
35	43	93.5	476	4	US-09-536-977-74	Sequence 74, Appl
36	43	93.5	479	2	US-08-037-816A-20	Sequence 20, Appl
37	43	93.5	479	2	US-08-037-816A-24	Sequence 24, Appl
38	43	93.5	479	2	US-08-530-146-20	Sequence 20, Appl
39	43	93.5	479	2	US-08-530-146-24	Sequence 24, Appl
40	43	93.5	479	5	PCT-US91-02166-10	Sequence 10, Appl
41	43	93.5	479	5	PCT-US91-02250-1	Sequence 1, Appli
42	43	93.5	483	3	US-08-889-841B-31	Sequence 31, Appl
43	43	93.5	483	4	US-09-419-362-31	Sequence 31, Appl
44	43	93.5	484	2	US-08-037-816A-22	Sequence 22, Appl
45	43	93.5	484	2	US-08-530-146-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-08-889-841B-36
; Sequence 36, Application US/08889841B
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE
; FILE REFERENCE: 14918-703CIP
; CURRENT APPLICATION NUMBER: US/08/889,841B
; CURRENT FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: US 60/676,737
; PRIOR FILING DATE: 1996-07-08
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 474
; TYPE: PRT
; ORGANISM: HIV
US-08-889-841B-36

Query Match 100.0%; Score 46; DB 3; Length 474;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        207 KPVVSTQLLL 216
```

RESULT 2
US-08-889-841B-39
; Sequence 39, Application US/08889841B
; GENERAL INFORMATION:
; APPLICANT: Berman, Phillip W.
; TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE
; FILE REFERENCE: 14918-703CIP
; CURRENT APPLICATION NUMBER: US/08/889,841B
; CURRENT FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: US 60/676,737
; PRIOR FILING DATE: 1996-07-08
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 39
; LENGTH: 474
; TYPE: PRT
; ORGANISM: HIV
US-08-889-841B-39

Query Match 100.0%; Score 46; DB 3; Length 474;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        207 KPVVSTQLLL 216


RESULT 3
US-09-419-362-36
; Sequence 36, Application US/09419362
; Patent No. 6585979
; GENERAL INFORMATION:
;  APPLICANT: Berman, Phillip W.
;  TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE
;  FILE REFERENCE: 14918-703DIV1
;  CURRENT APPLICATION NUMBER: US/09/419,362
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER: US 08/889,841
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: US 60/676,737
;  PRIOR FILING DATE: 1996-07-08
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: HIV
US-09-419-362-36

  Query Match             100.0%;  Score 46;  DB 4;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        207 KPVVSTQLLL 216


RESULT 4
US-09-419-362-39
; Sequence 39, Application US/09419362
; Patent No. 6585979
; GENERAL INFORMATION:
;  APPLICANT: Berman, Phillip W.
;  TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE
;  FILE REFERENCE: 14918-703DIV1
;  CURRENT APPLICATION NUMBER: US/09/419,362
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER: US 08/889,841
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: US 60/676,737
;  PRIOR FILING DATE: 1996-07-08
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 39
;   LENGTH: 474
;   TYPE: PRT
;   ORGANISM: HIV
US-09-419-362-39

  Query Match             100.0%;  Score 46;  DB 4;  Length 474;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        207 KPVVSTQLLL 216


RESULT 5
US-08-889-841B-2
; Sequence 2, Application US/08889841B
; GENERAL INFORMATION:
;  APPLICANT: Berman, Phillip W.
;  TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE
;  FILE REFERENCE: 14918-703CIP
;  CURRENT APPLICATION NUMBER: US/08/889,841B
;  CURRENT FILING DATE:  1997-07-08
;  PRIOR APPLICATION NUMBER: US 60/676,737
;  PRIOR FILING DATE: 1996-07-08
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 498
;   TYPE: PRT
;   ORGANISM: HIV
US-08-889-841B-2

  Query Match             100.0%;  Score 46;  DB 3;  Length 498;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        232 KPVVSTQLLL 241


RESULT 6
US-08-889-841B-5
; Sequence 5, Application US/08889841B
; GENERAL INFORMATION:
;  APPLICANT: Berman, Phillip W.
;  TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE
;  FILE REFERENCE: 14918-703CIP
;  CURRENT APPLICATION NUMBER: US/08/889,841B
;  CURRENT FILING DATE:  1997-07-08
;  PRIOR APPLICATION NUMBER: US 60/676,737
;  PRIOR FILING DATE: 1996-07-08
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 498
;   TYPE: PRT
;   ORGANISM: HIV
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(498)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-08-889-841B-5

  Query Match             100.0%;  Score 46;  DB 3;  Length 498;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        232 KPVVSTQLLL 241


RESULT 7
US-09-419-362-2
; Sequence 2, Application US/09419362
; Patent No. 6585979
; GENERAL INFORMATION:
;  APPLICANT: Berman, Phillip W.
;  TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE
;  FILE REFERENCE: 14918-703DIV1
;  CURRENT APPLICATION NUMBER: US/09/419,362
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER: US 08/889,841
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: US 60/676,737
;  PRIOR FILING DATE: 1996-07-08
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 498
```

```
;   TYPE: PRT
;   ORGANISM: HIV
US-09-419-362-2

  Query Match             100.0%;  Score 46;  DB 4;  Length 498;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        232 KPVVSTQLLL 241


RESULT 8
US-09-419-362-5
; Sequence 5, Application US/09419362
; Patent No. 6585979
; GENERAL INFORMATION:
;  APPLICANT: Berman, Phillip W.
;  TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE
;  FILE REFERENCE: 14918-703DIV1
;  CURRENT APPLICATION NUMBER: US/09/419,362
;  CURRENT FILING DATE:  1999-10-15
;  PRIOR APPLICATION NUMBER: US 08/889,841
;  PRIOR FILING DATE: 1997-07-08
;  PRIOR APPLICATION NUMBER: US 60/676,737
;  PRIOR FILING DATE: 1996-07-08
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 498
;   TYPE: PRT
;   ORGANISM: HIV
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(498)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-419-362-5

  Query Match             100.0%;  Score 46;  DB 4;  Length 498;
  Best Local Similarity   100.0%;  Pred. No. 0.14;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        232 KPVVSTQLLL 241


RESULT 9
US-09-646-028-50
; Sequence 50, Application US/09646028
; Patent No. 6562347
; GENERAL INFORMATION:
;  APPLICANT:  Kwak, Larry
;  APPLICANT:  Biragyn, Arya
;  TITLE OF INVENTION:  METHODS AND COMPOSITIONS OF
;  TITLE OF INVENTION:  CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
;  FILE REFERENCE:  14014.0316/P
;  CURRENT APPLICATION NUMBER: US/09/646,028
;  CURRENT FILING DATE:  2000-09-12
;  PRIOR APPLICATION NUMBER:  60/077,745
;  PRIOR FILING DATE:  1998-03-12
;  NUMBER OF SEQ ID NOS:  57
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 50
;   LENGTH: 587
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-09-646-028-50

  Query Match             100.0%;  Score 46;  DB 4;  Length 587;
  Best Local Similarity   100.0%;  Pred. No. 0.16;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        328 KPVVSTQLLL 337


RESULT 10
US-09-646-028-56
; Sequence 56, Application US/09646028
; Patent No. 6562347
; GENERAL INFORMATION:
;  APPLICANT:  Kwak, Larry
;  APPLICANT:  Biragyn, Arya
;  TITLE OF INVENTION:  METHODS AND COMPOSITIONS OF
;  TITLE OF INVENTION:  CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
;  FILE REFERENCE:  14014.0316/P
;  CURRENT APPLICATION NUMBER: US/09/646,028
;  CURRENT FILING DATE:  2000-09-12
;  PRIOR APPLICATION NUMBER:  60/077,745
;  PRIOR FILING DATE:  1998-03-12
;  NUMBER OF SEQ ID NOS:  57
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 56
;   LENGTH: 595
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-09-646-028-56

  Query Match             100.0%;  Score 46;  DB 4;  Length 595;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        336 KPVVSTQLLL 345


RESULT 11
US-08-392-806A-2
; Sequence 2, Application US/08392806A
; Patent No. 5965135
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  New HIV-1 virus isolates of a
;    TITLE OF INVENTION:  subtype, vaccine against HIV-1 virus infections of this subtype
;    TITLE OF INVENTION:  and method of producing same, use of the HIV-1 virus isolates
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jacobson, Price Holman and Stern, PLLC
;      STREET:  400 Seventh street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/392,806A
;      FILING DATE:  20-APR-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP93/02275
;      FILING DATE:  25-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P4228787.1
;      FILING DATE:  29-AUG-1992
```

```
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  596 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-392-806A-2

  Query Match             100.0%;  Score 46;  DB 2;  Length 596;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        137 KPVVSTQLLL 146


RESULT 12
US-09-257-490-2
; Sequence 2, Application US/09257490A
; Patent No. 6248328
; GENERAL INFORMATION:
;  APPLICANT: Dietrich, Ursula
;  APPLICANT: Von Briesen, Hagen
;  APPLICANT: Grez, Manuel
;  APPLICANT: Rubsamen-Waigmann, Helga
;  TITLE OF INVENTION: HIV-1 virus isolates of a subtype and its differential
;  TITLE OF INVENTION: diagnostics, a vaccine against HIV-1 virus infections
;  TITLE OF INVENTION: of this subtype and method of producing same, use of
;  TITLE OF INVENTION: the HIV-1 virus isolates
;  FILE REFERENCE: 10496/P58512US1
;  CURRENT APPLICATION NUMBER: US/09/257,490A
;  CURRENT FILING DATE: 1999-02-25
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 596
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus
;   FEATURE:
;   OTHER INFORMATION: HIV-1(D757)
US-09-257-490-2

  Query Match             100.0%;  Score 46;  DB 3;  Length 596;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        137 KPVVSTQLLL 146


RESULT 13
US-08-392-806A-4
; Sequence 4, Application US/08392806A
; Patent No. 5965135
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  New HIV-1 virus isolates of a
;    TITLE OF INVENTION:  subtype, vaccine against HIV-1 virus infections of this subtyp
;    TITLE OF INVENTION:  and method of producing same, use of the HIV-1 virus isolates
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jacobson, Price Holman and Stern, PLLC
;      STREET:  400 Seventh street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/392,806A
;      FILING DATE:  20-APR-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP93/02275
;      FILING DATE:  25-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P4228787.1
;      FILING DATE:  29-AUG-1992
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  600 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-392-806A-4

  Query Match             100.0%;  Score 46;  DB 2;  Length 600;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        139 KPVVSTQLLL 148


RESULT 14
US-09-257-490-4
; Sequence 4, Application US/09257490A
; Patent No. 6248328
; GENERAL INFORMATION:
;  APPLICANT: Dietrich, Ursula
;  APPLICANT: Von Briesen, Hagen
;  APPLICANT: Grez, Manuel
;  APPLICANT: Rubsamen-Waigmann, Helga
;  TITLE OF INVENTION: HIV-1 virus isolates of a subtype and its differential
;  TITLE OF INVENTION: diagnostics, a vaccine against HIV-1 virus infections
;  TITLE OF INVENTION: of this subtype and method of producing same, use of
;  TITLE OF INVENTION: the HIV-1 virus isolates
;  FILE REFERENCE: 10496/P58512US1
;  CURRENT APPLICATION NUMBER: US/09/257,490A
;  CURRENT FILING DATE: 1999-02-25
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 600
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus
;   FEATURE:
;   OTHER INFORMATION: HIV-1(D747)
US-09-257-490-4

  Query Match             100.0%;  Score 46;  DB 3;  Length 600;
  Best Local Similarity   100.0%;  Pred. No. 0.17;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        139 KPVVSTQLLL 148


RESULT 15
US-09-646-028-52
; Sequence 52, Application US/09646028
; Patent No. 6562347
; GENERAL INFORMATION:
;  APPLICANT:  Kwak, Larry
;  APPLICANT:  Biragyn, Arya
;  TITLE OF INVENTION:  METHODS AND COMPOSITIONS OF
;  TITLE OF INVENTION:  CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
;  FILE REFERENCE:  14014.0316/P
```

```
;  CURRENT APPLICATION NUMBER: US/09/646,028
;  CURRENT FILING DATE:  2000-09-12
;  PRIOR APPLICATION NUMBER:  60/077,745
;  PRIOR FILING DATE:  1998-03-12
;  NUMBER OF SEQ ID NOS:  57
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 52
;   LENGTH: 601
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-09-646-028-52

  Query Match           100.0%;  Score 46;  DB 4;  Length 601;
  Best Local Similarity  100.0%;  Pred. No. 0.17;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        342 KPVVSTQLLL 351


Search completed: March 12, 2004, 14:26:19
Job time : 23 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 12, 2004, 14:25:16 ; Search time 34 Seconds
(without alignments)
62.104 Million cell updates/sec

Title: PARKIN524.PEP
Perfect score: 46
Sequence: 1 kpvvstqlll 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	10	14	US-10-200-708-211	Sequence 211, App
2	46	100.0	10	14	US-10-200-708-328	Sequence 328, App
3	46	100.0	10	14	US-10-200-708-530	Sequence 530, App
4	46	100.0	236	10	US-09-827-688-2	Sequence 2, Appli
5	46	100.0	587	14	US-10-335-394-50	Sequence 50, Appl
6	46	100.0	595	14	US-10-335-394-56	Sequence 56, Appl
7	46	100.0	601	14	US-10-335-394-52	Sequence 52, Appl
8	46	100.0	803	14	US-10-190-435-134	Sequence 134, App
9	46	100.0	803	14	US-10-190-435-135	Sequence 135, App
10	46	100.0	845	14	US-10-190-435-129	Sequence 129, App
11	46	100.0	845	14	US-10-190-435-130	Sequence 130, App
12	46	100.0	845	14	US-10-190-435-143	Sequence 143, App
13	46	100.0	849	14	US-10-190-435-148	Sequence 148, App
14	46	100.0	851	14	US-10-190-435-131	Sequence 131, App
15	46	100.0	851	14	US-10-190-435-149	Sequence 149, App
16	46	100.0	853	14	US-10-190-435-133	Sequence 133, App
17	46	100.0	854	14	US-10-190-435-5	Sequence 5, Appli
18	46	100.0	854	14	US-10-190-435-128	Sequence 128, App
19	46	100.0	854	14	US-10-241-009-5	Sequence 5, Appli
20	46	100.0	854	14	US-10-190-434B-5	Sequence 5, Appli
21	46	100.0	854	14	US-10-190-305A-5	Sequence 5, Appli
22	46	100.0	855	14	US-10-190-435-144	Sequence 144, App
23	46	100.0	855	14	US-10-369-294-11	Sequence 11, Appl
24	46	100.0	855	14	US-10-369-294-12	Sequence 12, Appl
25	46	100.0	857	14	US-10-190-435-132	Sequence 132, App
26	46	100.0	857	14	US-10-190-435-138	Sequence 138, App
27	46	100.0	858	14	US-10-190-435-150	Sequence 150, App
28	46	100.0	859	14	US-10-190-435-137	Sequence 137, App
29	46	100.0	860	14	US-10-190-435-6	Sequence 6, Appli
30	46	100.0	860	14	US-10-190-435-145	Sequence 145, App
31	46	100.0	860	14	US-10-190-435-146	Sequence 146, App
32	46	100.0	860	14	US-10-241-009-6	Sequence 6, Appli
33	46	100.0	860	14	US-10-190-434B-6	Sequence 6, Appli
34	46	100.0	860	14	US-10-190-305A-6	Sequence 6, Appli
35	46	100.0	861	14	US-10-190-435-139	Sequence 139, App
36	46	100.0	862	14	US-10-190-435-141	Sequence 141, App
37	46	100.0	862	14	US-10-190-435-142	Sequence 142, App
38	46	100.0	865	14	US-10-190-435-140	Sequence 140, App
39	46	100.0	867	14	US-10-190-435-3	Sequence 3, Appli
40	46	100.0	867	14	US-10-190-435-126	Sequence 126, App
41	46	100.0	867	14	US-10-241-009-3	Sequence 3, Appli
42	46	100.0	867	14	US-10-190-434B-3	Sequence 3, Appli
43	46	100.0	867	14	US-10-190-305A-3	Sequence 3, Appli
44	46	100.0	869	14	US-10-190-435-4	Sequence 4, Appli
45	46	100.0	869	14	US-10-241-009-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-200-708-211
; Sequence 211, Application US/10200708
; Publication No. US20030180314A1
; GENERAL INFORMATION:
; APPLICANT: DeGroot, Anne S.
; TITLE OF INVENTION: HIV VACCINE CANDIDATE PEPTIDES
; FILE REFERENCE: 17999-001
; CURRENT APPLICATION NUMBER: US/10/200,708
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US/09/351,036
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/092,346
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: 60/115,145
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: 60/130,677
; PRIOR FILING DATE: 1999-04-23
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-10-200-708-211

Query Match 100.0%; Score 46; DB 14; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 1 KPVVSTQLLL 10

RESULT 2
US-10-200-708-328
; Sequence 328, Application US/10200708

```
; Publication No. US20030180314A1
; GENERAL INFORMATION:
;  APPLICANT: DeGroot, Anne S.
;  TITLE OF INVENTION: HIV VACCINE CANDIDATE PEPTIDES
;  FILE REFERENCE: 17999-001
;  CURRENT APPLICATION NUMBER: US/10/200,708
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/351,036
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 60/092,346
;  PRIOR FILING DATE: 1998-07-10
;  PRIOR APPLICATION NUMBER: 60/115,145
;  PRIOR FILING DATE: 1999-01-08
;  PRIOR APPLICATION NUMBER: 60/130,677
;  PRIOR FILING DATE: 1999-04-23
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 328
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus
US-10-200-708-328

  Query Match             100.0%;  Score 46;  DB 14;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 3
US-10-200-708-530
; Sequence 530, Application US/10200708
; Publication No. US20030180314A1
; GENERAL INFORMATION:
;  APPLICANT: DeGroot, Anne S.
;  TITLE OF INVENTION: HIV VACCINE CANDIDATE PEPTIDES
;  FILE REFERENCE: 17999-001
;  CURRENT APPLICATION NUMBER: US/10/200,708
;  CURRENT FILING DATE:  2002-07-22
;  PRIOR APPLICATION NUMBER: US/09/351,036
;  PRIOR FILING DATE: 1999-07-09
;  PRIOR APPLICATION NUMBER: 60/092,346
;  PRIOR FILING DATE: 1998-07-10
;  PRIOR APPLICATION NUMBER: 60/115,145
;  PRIOR FILING DATE: 1999-01-08
;  PRIOR APPLICATION NUMBER: 60/130,677
;  PRIOR FILING DATE: 1999-04-23
;  NUMBER OF SEQ ID NOS: 672
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 530
;   LENGTH: 10
;   TYPE: PRT
;   ORGANISM: Human immunodeficiency virus
US-10-200-708-530

  Query Match             100.0%;  Score 46;  DB 14;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          1 KPVVSTQLLL 10


RESULT 4
US-09-827-688-2
; Sequence 2, Application US/09827688
; Publication No. US20030165476A1
; GENERAL INFORMATION:
;  APPLICANT: ORSON, FRANK
;  APPLICANT:  KINSEY, BERMA
;  APPLICANT:  BHOGAL, BALBIR
;  TITLE OF INVENTION: MACROAGGREGATED PROTEIN CONJUGATES AS ORAL GENETIC IMMUNIZATION DE
;  TITLE OF INVENTION:  AGENTS
;  FILE REFERENCE: P01949US1/10004014
;  CURRENT APPLICATION NUMBER: US/09/827,688
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: 60/195,680
;  PRIOR FILING DATE: 2000-04-07
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
;   LENGTH: 236
;   TYPE: PRT
;   ORGANISM: HIV U#23
US-09-827-688-2

  Query Match             100.0%;  Score 46;  DB 10;  Length 236;
  Best Local Similarity   100.0%;  Pred. No. 0.31;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db         46 KPVVSTQLLL 55


RESULT 5
US-10-335-394-50
; Sequence 50, Application US/10335394
; Publication No. US20030138452A1
; GENERAL INFORMATION:
;  APPLICANT: Kwak, Larry
;  APPLICANT:  Biragyn, Arya
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
;  TITLE OF INVENTION:  CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
;  FILE REFERENCE: 14014.0316/P
;  CURRENT APPLICATION NUMBER: US/10/335,394
;  CURRENT FILING DATE:  2002-12-31
;  PRIOR APPLICATION NUMBER: US/09/646,028
;  PRIOR FILING DATE: 2000-09-12
;  PRIOR APPLICATION NUMBER: 60/077,745
;  PRIOR FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 50
;   LENGTH: 587
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-10-335-394-50

  Query Match             100.0%;  Score 46;  DB 14;  Length 587;
  Best Local Similarity   100.0%;  Pred. No. 0.84;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        328 KPVVSTQLLL 337


RESULT 6
US-10-335-394-56
; Sequence 56, Application US/10335394
; Publication No. US20030138452A1
; GENERAL INFORMATION:
;  APPLICANT: Kwak, Larry
;  APPLICANT:  Biragyn, Arya
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
;  TITLE OF INVENTION:  CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
;  FILE REFERENCE: 14014.0316/P
;  CURRENT APPLICATION NUMBER: US/10/335,394
;  CURRENT FILING DATE:  2002-12-31
```

```
;  PRIOR APPLICATION NUMBER: US/09/646,028
;  PRIOR FILING DATE: 2000-09-12
;  PRIOR APPLICATION NUMBER: 60/077,745
;  PRIOR FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 56
;   LENGTH: 595
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-10-335-394-56

  Query Match             100.0%;  Score 46;  DB 14;  Length 595;
  Best Local Similarity   100.0%;  Pred. No. 0.85;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        336 KPVVSTQLLL 345


RESULT 7
US-10-335-394-52
; Sequence 52, Application US/10335394
; Publication No. US20030138452A1
; GENERAL INFORMATION:
;  APPLICANT: Kwak, Larry
;  APPLICANT:  Biragyn, Arya
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS OF
;  TITLE OF INVENTION:  CHEMOKINE-TUMOR ANTIGEN FUSION PROTEINS AS CANCER VACCINES
;  FILE REFERENCE: 14014.0316/P
;  CURRENT APPLICATION NUMBER: US/10/335,394
;  CURRENT FILING DATE:  2002-12-31
;  PRIOR APPLICATION NUMBER: US/09/646,028
;  PRIOR FILING DATE: 2000-09-12
;  PRIOR APPLICATION NUMBER: 60/077,745
;  PRIOR FILING DATE: 1998-03-12
;  NUMBER OF SEQ ID NOS: 57
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 52
;   LENGTH: 601
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of artificial sequence:/note=synthetic construct
US-10-335-394-52

  Query Match             100.0%;  Score 46;  DB 14;  Length 601;
  Best Local Similarity   100.0%;  Pred. No. 0.86;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        342 KPVVSTQLLL 351


RESULT 8
US-10-190-435-134
; Sequence 134, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT:  BARNETT, Susan W.
;  APPLICANT:  LIAN, Ying
;  APPLICANT:  ENGELBRECHT, Susan
;  APPLICANT:  VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION:  POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE:  2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 134
;   LENGTH: 803
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: TV007cB104
US-10-190-435-134

  Query Match             100.0%;  Score 46;  DB 14;  Length 803;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        252 KPVVSTQLLL 261


RESULT 9
US-10-190-435-135
; Sequence 135, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT:  BARNETT, Susan W.
;  APPLICANT:  LIAN, Ying
;  APPLICANT:  ENGELBRECHT, Susan
;  APPLICANT:  VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION:  POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE:  2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 135
;   LENGTH: 803
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: TV007cB105
US-10-190-435-135

  Query Match             100.0%;  Score 46;  DB 14;  Length 803;
  Best Local Similarity   100.0%;  Pred. No. 1.2;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        252 KPVVSTQLLL 261


RESULT 10
US-10-190-435-129
; Sequence 129, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT:  BARNETT, Susan W.
;  APPLICANT:  LIAN, Ying
;  APPLICANT:  ENGELBRECHT, Susan
;  APPLICANT:  VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION:  POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE:  2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
;   LENGTH: 845
;   TYPE: PRT
```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TV012c2.1
US-10-190-435-129

Query Match 100.0%; Score 46; DB 14; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 242 KPVVSTQLLL 251

RESULT 11
US-10-190-435-130
; Sequence 130, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 130
; LENGTH: 845
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TV012c2.2
US-10-190-435-130

Query Match 100.0%; Score 46; DB 14; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 242 KPVVSTQLLL 251

RESULT 12
US-10-190-435-143
; Sequence 143, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 143
; LENGTH: 845
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TV003cE260
US-10-190-435-143

Query Match 100.0%; Score 46; DB 14; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 240 KPVVSTQLLL 249

RESULT 13
US-10-190-435-148
; Sequence 148, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 148
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 96BW05.02
US-10-190-435-148

Query Match 100.0%; Score 46; DB 14; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 250 KPVVSTQLLL 259

RESULT 14
US-10-190-435-131
; Sequence 131, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
; APPLICANT: ZUR MEGEDE, Jan
; APPLICANT: BARNETT, Susan W.
; APPLICANT: LIAN, Ying
; APPLICANT: ENGELBRECHT, Susan
; APPLICANT: VAN RENSBURG, Estrelita J.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
; TITLE OF INVENTION: POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: PP18133.003 / 2302-18133
; CURRENT APPLICATION NUMBER: US/10/190,435
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 319
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 131
; LENGTH: 851
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TV006c9.1
US-10-190-435-131

Query Match 100.0%; Score 46; DB 14; Length 851;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10

```
                  ||||||||||
Db        247 KPVVSTQLLL 256


RESULT 15
US-10-190-435-149
; Sequence 149, Application US/10190435
; Publication No. US20030143248A1
; GENERAL INFORMATION:
;  APPLICANT: ZUR MEGEDE, Jan
;  APPLICANT:  BARNETT, Susan W.
;  APPLICANT:  LIAN, Ying
;  APPLICANT:  ENGELBRECHT, Susan
;  APPLICANT:  VAN RENSBURG, Estrelita J.
;  TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
;  TITLE OF INVENTION:  POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF
;  FILE REFERENCE: PP18133.003 / 2302-18133
;  CURRENT APPLICATION NUMBER: US/10/190,435
;  CURRENT FILING DATE:  2002-12-30
;  NUMBER OF SEQ ID NOS: 319
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 149
;   LENGTH: 851
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: ETH2220
US-10-190-435-149

  Query Match            100.0%;  Score 46;  DB 14;  Length 851;
  Best Local Similarity  100.0%;  Pred. No. 1.3;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        249 KPVVSTQLLL 258


Search completed: March 12, 2004, 14:30:42
Job time : 34 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 12, 2004, 14:22:50 ; Search time 21 Seconds
(without alignments)
45.805 Million cell updates/sec

Title: PARKIN524.PEP
Perfect score: 46
Sequence: 1 kpvvstqlll 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	495	2	S31493	env polyprotein -
2	46	100.0	859	2	T01672	envelope polyprote
3	43	93.5	219	2	S25939	env protein - huma
4	43	93.5	729	1	VCLJKX	env polyprotein pr
5	43	93.5	843	1	H44001	env polyprotein pr
6	43	93.5	846	1	VCLJND	env polyprotein pr
7	43	93.5	847	2	T09448	envelope glycoprot
8	43	93.5	847	2	S13289	env protein - huma
9	43	93.5	851	2	S33985	env polyprotein -
10	43	93.5	852	1	VCLJBR	env polyprotein -
11	43	93.5	852	2	T12016	envelope glycoprot
12	43	93.5	853	2	S54384	envelope polyprote
13	43	93.5	854	2	S13288	env protein - huma
14	43	93.5	855	1	VCLJZR	env polyprotein pr
15	43	93.5	856	1	VCLJH3	env polyprotein pr
16	43	93.5	856	1	VCLJVL	env polyprotein pr
17	43	93.5	856	1	VCLJ3W	env polyprotein pr
18	43	93.5	856	1	A44963	env polyprotein pr
19	43	93.5	859	1	VCLJMN	env polyprotein pr
20	43	93.5	861	1	VCLJLV	env polyprotein pr
21	43	93.5	861	1	VCLJKB	env polyprotein pr
22	43	93.5	868	1	VCLJH4	env polyprotein -
23	42	91.3	506	2	A40218	envelop glycoprote
24	42	91.3	855	1	VCLJA2	env polyprotein pr
25	41	89.1	854	1	VCLJSI	env polyprotein pr
26	40	87.0	877	2	S49197	envelope protein p
27	38	82.6	861	1	VCLJSC	env polyprotein pr
28	37	80.4	863	2	A53034	gag polyprotein -
29	34	73.9	224	2	S71749	DCL protein precur
30	34	73.9	443	2	C41621	env polyprotein P
31	34	73.9	445	2	A41621	env polyprotein M
32	34	73.9	454	2	B41621	env polyprotein D
33	34	73.9	486	2	S06949	achaete-scute locu
34	33	71.7	540	2	T51390	TCP-1 chaperonin-l
35	33	71.7	876	2	AB1177	cation transportin
36	33	71.7	876	2	AF1534	cation transportin
37	32	69.6	130	2	T51684	probable transcrip
38	32	69.6	213	2	F84532	hypothetical prote
39	32	69.6	254	2	C70387	dehydrogenase - Aq
40	32	69.6	255	2	S50326	phosphoglycerate m
41	32	69.6	369	2	I64234	competence locus E
42	32	69.6	454	2	S61019	hypothetical prote
43	32	69.6	638	2	T22518	hypothetical prote
44	32	69.6	662	2	G82721	regulator of patho
45	32	69.6	868	2	B82213	aconitate hydratas

ALIGNMENTS

RESULT 1
S31493
env polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-Aug-1999
C;Accession: S31493
R;Clegg, C.S.
submitted to the EMBL Data Library, December 1992
A;Reference number: S31493
A;Accession: S31493
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-495 <CLE>
A;Cross-references: EMBL:Z19533; NID:g60244; PIDN:CAA79593.1; PID:g60245
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: polyprotein

Query Match 100.0%; Score 46; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 233 KPVVSTQLLL 242

RESULT 2
T01672
envelope polyprotein precursor - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 21-Jul-2000
C;Accession: T01672
R;Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.
Cell 46, 63-74, 1986
A;Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isola
A;Reference number: Z14389; MUID:86245056; PMID:2424612
A;Accession: T01672
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-859 <ALI>
A;Cross-references: EMBL:K03456; NID:g60228; PIDN:CAA28016.1; PID:g60234
C;Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 46; DB 2; Length 859;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||
Db 257 KPVVSTQLLL 266

RESULT 3
S25939
env protein - human immunodeficiency virus type 1 (fragment)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 25-Feb-1994 #sequence_revision 30-Jan-1998 #text_change 26-Aug-1999
C;Accession: S25939
R;Guo, H.G.; Chermann, J.C.; Waters, D.; Hall, L.; Louie, A.; Gallo, R.C.; Streicher, H.
Nature 349, 745-746, 1991
A;Title: Sequence analysis of original HIV-1.
A;Reference number: S25937; MUID:91156044; PMID:2000145
A;Accession: S25939
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-219 <GUO>
A;Cross-references: EMBL:X57447; NID:g60212; PIDN:CAA40693.1; PID:g388536
A;Experimental source: strain JBB
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1991
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: coat protein

Query Match 93.5%; Score 43; DB 2; Length 219;
Best Local Similarity 90.0%; Pred. No. 0.13;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
:|||||||||
Db 209 RPVVSTQLLL 218

RESULT 4
VCLJKX
env polyprotein precursor - human immunodeficiency virus type 1 (strain KB-1-gp32)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp32
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 01-Mar-1996
C;Accession: B42995
R;Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.
Virology 189, 534-546, 1992
A;Title: Analysis of a human immunodeficiency virus type 1 isolate carrying a truncated
A;Reference number: A42995; MUID:92351552; PMID:1322587
A;Accession: B42995
A;Molecule type: mRNA
A;Residues: 1-729 <SHI>
A;Cross-references: GB:S41266; GB:D01206
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-689/Domain: extracellular #status predicted <EXT>
F;1-33/Domain: signal sequence #status predicted <SIG>
F;17-33/Region: hydrophobic #status predicted
F;34-517/Product: coat protein gp120 #status predicted <CP1>
F;514-517/Region: cleavage processing #status predicted
F;518-729/Product: coat protein gp32 #status predicted <CP2>
F;518-534/Region: hydrophobic #status predicted
F;690-711/Domain: transmembrane #status predicted <TM1>
F;712-729/Domain: intracellular #status predicted <INT>
F;93,141,145,146,163,191,192,237,241,248,269,283,296,308,338,345,361,367,397,403,408,414

Query Match 93.5%; Score 43; DB 1; Length 729;
Best Local Similarity 90.0%; Pred. No. 0.51;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
:|||||||||
Db 259 RPVVSTQLLL 268

RESULT 5
H44001
env polyprotein precursor - human immunodeficiency virus type 1 (strain YU-2)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994
C;Accession: H44001
R;Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.
J. Virol. 66, 6587-6600, 1992
A;Title: Complete nucleotide sequence, genome organization, and biological properties o
A;Reference number: A44001; MUID:93021387; PMID:1404605
A;Accession: H44001
A;Molecule type: DNA
A;Residues: 1-843 <LIY>
A;Cross-references: GB:M93258
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;19-35/Region: hydrophobic
F;30-489/Product: coat protein gp120 #status predicted <GP1>
F;490-843/Product: coat protein gp41 #status predicted <GP2>
F;499-515/Region: hydrophobic
F;673-689/Region: hydrophobic
F;738-755/Domain: transmembrane #status predicted <TMN>
F;87,129,135,138,154,158,184,193,230,237,258,272,285,291,297,327,351,381,389,395,400,435

Query Match 93.5%; Score 43; DB 1; Length 843;
Best Local Similarity 90.0%; Pred. No. 0.6;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
:|||||||||
Db 248 RPVVSTQLLL 257

RESULT 6
VCLJND
env polyprotein precursor - human immunodeficiency virus type 1 (isolate NDK)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C;Accession: JQ0066
R;Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;
Gene 81, 275-284, 1989
A;Title: Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human immunoc
A;Reference number: JQ0065; MUID:90034200; PMID:2806917
A;Accession: JQ0066
A;Molecule type: DNA
A;Residues: 1-846 <SPI>
A;Cross-references: GB:M27323; NID:g328154; PIDN:AAA44873.1; PID:g328162
A;Note: the authors translated the codon GCG for residue 523 as Arg
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprote
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-501/Product: coat protein gp120 #status predicted <CP1>
F;502-846/Product: coat protein gp41 #status predicted <CP2>
F;502-520/Domain: transmembrane #status predicted <TM1>
F;674-692/Domain: transmembrane #status predicted <TM2>
F;87,129,151,179,182,229,236,257,271,284,290,351,382,388,392,395,401,438,451,452,601,606,

Query Match 93.5%; Score 43; DB 1; Length 846;
Best Local Similarity 90.0%; Pred. No. 0.6;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
:|||||||||

```
Db        247 RPVVSTQLLL 256


RESULT 7
T09448
envelope glycoprotein - human immunodeficiency virus type 1 (strain JRFL)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C;Accession: T09448
R;Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,
submitted to the EMBL Data Library, July 1996
A;Reference number: Z16673
A;Accession: T09448
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-847 <PAN>
A;Cross-references: EMBL:U63632; NID:g1465777; PID:g1465781
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein

  Query Match             93.5%;  Score 43;  DB 2;  Length 847;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        249 RPVVSTQLLL 258


RESULT 8
S13289
env protein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C;Accession: S13289
R;O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A
Nature 348, 69-73, 1990
A;Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A;Reference number: S13288; MUID:91043044; PMID:2172833
A;Accession: S13289
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-847 <OBR>
C;Superfamily: type E retrovirus env polyprotein

  Query Match             93.5%;  Score 43;  DB 2;  Length 847;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        249 RPVVSTQLLL 258


RESULT 9
S33985
env polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 26-Aug-1999
C;Accession: S33985
R;Carlini, F.
submitted to the EMBL Data Library, November 1991
A;Reference number: S33979
A;Accession: S33985
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-851 <CAR>
A;Cross-references: EMBL:Z11530; NID:g60192; PIDN:CAA77628.1; PID:g60199
C;Superfamily: type E retrovirus env polyprotein

  Query Match             93.5%;  Score 43;  DB 2;  Length 851;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        252 RPVVSTQLLL 261


RESULT 10
VCLJBR
env polyprotein - human immunodeficiency virus type 1 (isolate BR)
N;Alternate names: coat polyprotein
N;Contains: coat protein gp120; coat protein gp41
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Feb-1997
C;Accession: A31667
R;Anand, R.; Thayer, R.; Srinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar, S.
Virology 168, 79-89, 1989
A;Title: Biological and molecular characterization of human immunodeficiency virus (HIV-1
A;Reference number: A94389; MUID:89085613; PMID:2789516
A;Accession: A31667
A;Molecule type: DNA
A;Residues: 1-852 <ANA>
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: capsid protein; coat protein; polyprotein; transmembrane protein
F;1-516/Product: coat protein gp120 #status predicted <CP1>
F;517-852/Product: coat protein gp41 #status predicted <CP2>

  Query Match             93.5%;  Score 43;  DB 1;  Length 852;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        252 RPVVSTQLLL 261


RESULT 11
T12016
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C;Accession: T12016
R;McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.
AIDS Res. Hum. Retroviruses 14, 329-337, 1998
A;Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in Sa
A;Reference number: Z17379; MUID:98178716; PMID:9519894
A;Accession: T12016
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-852 <MCC>
A;Cross-references: EMBL:U90934; NID:g2351783; PIDN:AAC59271.1; PID:g2351784
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein

  Query Match             93.5%;  Score 43;  DB 2;  Length 852;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        251 RPVVSTQLLL 260


RESULT 12
S54384
envelope polyprotein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 26-Aug-1999
C;Accession: S54384
R;Theodore, T.; Buckler-White, A.J.
submitted to the EMBL Data Library, July 1989
```

```
A;Reference number: S54377
A;Accession: S54384
A;Status: preliminary
A;Molecule type: genomic RNA
A;Residues: 1-853 <THE>
A;Cross-references: EMBL:M22639; NID:g329377; PIDN:AAA45370.1; PID:g329385
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: polyprotein

  Query Match             93.5%;  Score 43;  DB 2;  Length 853;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        253 RPVVSTQLLL 262


RESULT 13
S13288
env protein - human immunodeficiency virus type 1
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C;Accession: S13288
R;O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A
Nature 348, 69-73, 1990
A;Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120
A;Reference number: S13288; MUID:91043044; PMID:2172833
A;Accession: S13288
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-854 <OBR>
C;Superfamily: type E retrovirus env polyprotein

  Query Match             93.5%;  Score 43;  DB 2;  Length 854;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        250 RPVVSTQLLL 259


RESULT 14
VCLJZR
env polyprotein precursor - human immunodeficiency virus Zr-6
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus Zr-6
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: D26192
R;Srinivasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorino, P.; Schochetman, G.; Cu
Gene 52, 71-82, 1987
A;Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleoti
A;Reference number: A26192; MUID:87248097; PMID:3036660
A;Accession: D26192
A;Molecule type: DNA
A;Residues: 1-855 <SRI>
A;Cross-references: GB:K03458; GB:M16322; NID:g329398; PIDN:AAA45380.1; PID:g329403
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-855/Product: env polyprotein #status predicted <MAT>
F;20-500/Product: exterior membrane glycoprotein #status predicted <EXT>
F;501-855/Product: transmembrane glycoprotein #status predicted <TMM>
F;87,129,140,145,154,158,186,189,199,236,243,264,278,291,297,333,340,355,386,392,398,404

  Query Match             93.5%;  Score 43;  DB 1;  Length 855;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        254 RPVVSTQLLL 263


RESULT 15
VCLJH3
env polyprotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 16-Jul-1999
C;Accession: A03973
R;Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dorar
nberger, J.A.; Papas, T.S.; Ghrayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.
Nature 313, 277-284, 1985
A;Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.
A;Reference number: A93353; MUID:85111123; PMID:2578615
A;Accession: A03973
A;Molecule type: DNA
A;Residues: 1-856 <RAT>
A;Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:g326383; PIDN:AAA442(
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprote
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-511/Product: exterior membrane glycoprotein #status predicted <EXT>
F;512-856/Product: transmembrane glycoprotein #status predicted <TMM>
F;88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406,
F;611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predict

  Query Match             93.5%;  Score 43;  DB 1;  Length 856;
  Best Local Similarity   90.0%;  Pred. No. 0.61;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        252 RPVVSTQLLL 261


Search completed: March 12, 2004, 14:25:44
Job time : 22 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 12, 2004, 14:22:50 ; Search time 11 Seconds
(without alignments)
47.337 Million cell updates/sec

Title: PARKIN524.PEP
Perfect score: 46
Sequence: 1 kpvvstqlll 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	421	1	ENV_HV1N5	P12490 human immun
2	46	100.0	852	1	ENV_HV1S3	P19549 human immun
3	46	100.0	855	1	ENV_HV1OY	P20888 human immun
4	46	100.0	856	1	ENV_HV1LW	Q70626 human immun
5	46	100.0	859	1	ENV_HV1MA	P04583 human immun
6	46	100.0	867	1	ENV_HV1J3	P12489 human immun
7	43	93.5	460	1	ENV_HV1Z3	P12491 human immun
8	43	93.5	843	1	ENV_HV1Y2	P35961 human immun
9	43	93.5	846	1	ENV_HV1ND	P18799 human immun
10	43	93.5	847	1	ENV_HV1S1	P19550 human immun
11	43	93.5	847	1	ENV_HV1W2	P05880 human immun
12	43	93.5	848	1	ENV_HV1JR	P20871 human immun
13	43	93.5	851	1	ENV_HV1B8	P04582 human immun
14	43	93.5	852	1	ENV_HV1BN	P12488 human immun
15	43	93.5	853	1	ENV_HV1EL	P04581 human immun
16	43	93.5	853	1	ENV_HV1MF	P19551 human immun
17	43	93.5	853	1	ENV_HV1Z2	P12487 human immun
18	43	93.5	855	1	ENV_HV1Z6	P04580 human immun
19	43	93.5	856	1	ENV_HV1B1	P03375 human immun
20	43	93.5	856	1	ENV_HV1H2	P04578 human immun
21	43	93.5	856	1	ENV_HV1H3	P04624 human immun
22	43	93.5	856	1	ENV_HV1MN	P05877 human immun
23	43	93.5	856	1	ENV_HV1PV	P03376 human immun
24	43	93.5	856	1	ENV_HV1W1	P31872 human immun
25	43	93.5	856	1	ENV_HV1ZH	P05881 human immun
26	43	93.5	861	1	ENV_HV1BR	P03377 human immun
27	43	93.5	861	1	ENV_HV1KB	P31819 human immun
28	43	93.5	863	1	ENV_HV1Z8	P05882 human immun
29	43	93.5	865	1	ENV_HV1RH	P04579 human immun
30	43	93.5	868	1	ENV_HV1C4	P05879 human immun
31	42	91.3	855	1	ENV_HV1A2	P03378 human immun
32	41	89.1	854	1	ENV_SIVCZ	P17281 chimpanzee
33	38	82.6	856	1	ENV_HV1SC	P05878 human immun
34	34	73.9	224	1	DCL_LYCES	Q42463 lycopersico
35	34	73.9	486	1	AST8_DROME	P09775 drosophila
36	32	69.6	369	1	Y316_MYCGE	P47558 mycoplasma
37	32	69.6	638	1	YT32_CAEEL	Q20680 caenorhabdi
38	32	69.6	1317	1	RPOD_SYNY3	P73334 synechocyst
39	31.5	68.5	843	1	NUOG_STRCO	Q9xar0 streptomyce
40	31	67.4	114	1	CO7R_HUMAN	O14548 homo sapien
41	31	67.4	264	1	Y736_SHEON	P46149 shewanella
42	31	67.4	417	1	YNFM_ECOLI	P43531 escherichia
43	31	67.4	459	1	TRME_BACSU	P25811 bacillus su
44	31	67.4	484	1	PEN3_ADEP3	Q84176 porcine ade
45	31	67.4	531	1	UGTE_CAEEL	Q10941 caenorhabdi

ALIGNMENTS

```
RESULT 1
ENV_HV1N5
ID   ENV_HV1N5       STANDARD;      PRT;   421 AA.
AC   P12490;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11698;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86259728; PubMed=3014529;
RA   Willey R.W., Rutledge R.A., Dias S., Folks T., Theodore T.,
RA   Buckler C.E., Martin M.A.;
RT   "Identification of conserved and divergent domains within the
RT   envelope gene of the acquired immunodeficiency syndrome retrovirus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5038-5042(1986).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03346; AAB02407.1; -.
DR   HIV; K03346; ENV$NY5.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL        1     30
FT   CHAIN        31    421       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   DISULFID     53     73       BY SIMILARITY.
FT   DISULFID    118    200       BY SIMILARITY.
FT   DISULFID    125    191       BY SIMILARITY.
FT   DISULFID    130    152       BY SIMILARITY.
FT   DISULFID    213    242       BY SIMILARITY.
FT   DISULFID    223    234       BY SIMILARITY.
FT   DISULFID    291    325       BY SIMILARITY.
FT   DISULFID    378    410       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     97     97       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    135    135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    140    140       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    155    155       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    183    183       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    229    229       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD     236    236       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     257    257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     271    271       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     284    284       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     290    290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     296    296       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     326    326       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     333    333       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     349    349       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     355    355       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     385    385       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     391    391       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     395    395       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   421 AA;  47493 MW;  25A575719C22967B CRC64;

  Query Match             100.0%;  Score 46;  DB 1;  Length 421;
  Best Local Similarity   100.0%;  Pred. No. 0.052;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        247 KPVVSTQLLL 256


RESULT 2
ENV_HV1S3
ID   ENV_HV1S3      STANDARD;      PRT;   852 AA.
AC   P19549;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11690;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90317906; PubMed=2370688;
RA   York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT   "Human immunodeficiency virus type 1 cellular host range,
RT   replication, and cytopathicity are linked to the envelope region of
RT   the viral genome.";
RL   J. Virol. 64:4016-4020(1990).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AY352275; AAQ17031.1; -.
DR   PDB; 1MEQ; 11-DEC-02.
DR   HIV; M38427; ENV$SF33.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW   3D-structure.
FT   SIGNAL        1     31       BY SIMILARITY.
FT   CHAIN        32    506       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN       507    852       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID     53     73       BY SIMILARITY.
FT   DISULFID    118    206       BY SIMILARITY.
FT   DISULFID    125    197       BY SIMILARITY.
FT   DISULFID    130    156       BY SIMILARITY.
FT   DISULFID    219    248       BY SIMILARITY.
FT   DISULFID    229    240       BY SIMILARITY.
FT   DISULFID    297    331       BY SIMILARITY.
FT   DISULFID    377    439       BY SIMILARITY.
FT   DISULFID    384    412       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    136    136       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    141    141       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    142    142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    155    155       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    159    159       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    189    189       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    198    198       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    290    290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    296    296       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    332    332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    355    355       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    385    385       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    391    391       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    397    397       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    401    401       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    405    405       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    442    442       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    457    457       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    607    607       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    612    612       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    621    621       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    633    633       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    812    812       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   852 AA;  96663 MW;  EE7BBF8D23C9910D CRC64;

  Query Match             100.0%;  Score 46;  DB 1;  Length 852;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        253 KPVVSTQLLL 262


RESULT 3
ENV_HV1OY
ID   ENV_HV1OY      STANDARD;      PRT;   855 AA.
AC   P20888;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (OYI isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11699;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90148544; PubMed=2559749;
RA   Huet T., Dazza M.C., Brun-Vezinet F., Roelants G.E., Wain-Hobson S.;
RT   "A highly defective HIV-1 strain isolated from a healthy Gabonese
RT   individual presenting an atypical western blot.";
RL   AIDS 3:707-715(1989).
CC   -!- MISCELLANEOUS: THE OYI ISOLATE WAS TAKEN FROM THE BLOOD OF A
CC       HEALTHY GABONESE INDIVIDUAL.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; M26727; AAA83397.1; -.
DR   HIV; M26727; ENV$OYI.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL        1     29
FT   CHAIN        30    509       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN       510    855       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID     53     73       BY SIMILARITY.
FT   DISULFID    118    210       BY SIMILARITY.
FT   DISULFID    125    201       BY SIMILARITY.
FT   DISULFID    130    162       BY SIMILARITY.
FT   DISULFID    223    252       BY SIMILARITY.
FT   DISULFID    233    244       BY SIMILARITY.
FT   DISULFID    301    335       BY SIMILARITY.
FT   DISULFID    381    442       BY SIMILARITY.
FT   DISULFID    388    415       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    134    134       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    142    142       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    145    145       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    165    165       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    202    202       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    267    267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    281    281       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    300    300       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    306    306       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    359    359       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    389    389       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    395    395       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    399    399       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    405    405       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    458    458       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    610    610       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    615    615       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    624    624       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    636    636       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    815    815       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   855 AA;  97476 MW;  9CF82A607ADD62DA CRC64;

  Query Match             100.0%;  Score 46;  DB 1;  Length 855;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        257 KPVVSTQLLL 266


RESULT 4
ENV_HV1LW
ID   ENV_HV1LW      STANDARD;      PRT;   856 AA.
AC   Q70626;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (LW12.3 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=82834;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95127297; PubMed=7826699;
RA   Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
RA   Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
RT   "Viral variability and serum antibody response in a laboratory worker
RT   infected with HIV type 1 (HTLV type IIIB).";
RL   AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U12055; AAA76690.1; -.
DR   PDB; 1IF3; 02-MAY-01.
DR   GlycoSuiteDB; Q70626; -.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW   3D-structure.
FT   SIGNAL        1     30
FT   CHAIN        31    511       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN       512    856       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID     54     74       BY SIMILARITY.
FT   DISULFID    119    205       BY SIMILARITY.
FT   DISULFID    126    196       BY SIMILARITY.
FT   DISULFID    131    157       BY SIMILARITY.
FT   DISULFID    218    247       BY SIMILARITY.
FT   DISULFID    228    239       BY SIMILARITY.
FT   DISULFID    296    331       BY SIMILARITY.
FT   DISULFID    378    445       BY SIMILARITY.
FT   DISULFID    385    418       BY SIMILARITY.
FT   CARBOHYD     88     88       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    136    136       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    141    141       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    156    156       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    160    160       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    230    230       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    241    241       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    262    262       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    276    276       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    289    289       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    301    301       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    332    332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    339    339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    356    356       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    386    386       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    392    392       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    397    397       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    406    406       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    448    448       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    463    463       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    611    611       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    616    616       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    624    624       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    637    637       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    674    674       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    750    750       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    816    816       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   856 AA;  96938 MW;  0C241332CF7E6687 CRC64;

  Query Match             100.0%;  Score 46;  DB 1;  Length 856;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        252 KPVVSTQLLL 261


RESULT 5
ENV_HV1MA
ID   ENV_HV1MA       STANDARD;      PRT;   859 AA.
AC   P04583;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11697;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86245056; PubMed=2424612;
RA   Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
RT   "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT   of two isolates from African patients.";
RL   Cell 46:63-74(1986).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04415; CAA28016.1; -.
DR   EMBL; A07116; CAA00623.1; -.
DR   PIR; T01672; T01672.
DR   HIV; K03456; ENV$MAL.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL       1     30
FT   CHAIN       31    513       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN      514    859       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID    53     73       BY SIMILARITY.
FT   DISULFID   118    210       BY SIMILARITY.
FT   DISULFID   125    201       BY SIMILARITY.
FT   DISULFID   130    162       BY SIMILARITY.
FT   DISULFID   223    252       BY SIMILARITY.
FT   DISULFID   233    244       BY SIMILARITY.
FT   DISULFID   301    334       BY SIMILARITY.
FT   DISULFID   380    445       BY SIMILARITY.
FT   DISULFID   387    418       BY SIMILARITY.
FT   CARBOHYD    87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   134    134       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   139    139       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   146    146       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   161    161       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   193    193       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   202    202       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   267    267       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   281    281       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   294    294       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   300    300       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   306    306       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   337    337       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   357    357       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   364    364       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   388    388       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   394    394       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   406    406       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   448    448       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   460    460       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   467    467       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   613    613       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   618    618       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   627    627       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   639    639       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   819    819       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   859 AA;  97109 MW;  DBCF9AA52E3ABF29 CRC64;

  Query Match             100.0%;  Score 46;  DB 1;  Length 859;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        257 KPVVSTQLLL 266


RESULT 6
ENV_HV1J3
ID   ENV_HV1J3       STANDARD;      PRT;   867 AA.
AC   P12489;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11694;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89352108; PubMed=2669897;
RA   Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;
RT   "Nucleotide sequences of gag and env genes of a Japanese isolate of
RT   HIV-1 and their expression in bacteria.";
RL   AIDS Res. Hum. Retroviruses 5:411-419(1989).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21138; AAB03526.1; -.
DR   HIV; M21138; ENV$JH3.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL       1     30
FT   CHAIN       31    516       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN      517    867       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID    53     73       BY SIMILARITY.
FT   DISULFID   118    217       BY SIMILARITY.
FT   DISULFID   125    208       BY SIMILARITY.
FT   DISULFID   130    160       BY SIMILARITY.
FT   DISULFID   230    259       BY SIMILARITY.
FT   DISULFID   240    251       BY SIMILARITY.
FT   DISULFID   308    342       BY SIMILARITY.
FT   DISULFID   388    457       BY SIMILARITY.
FT   DISULFID   395    430       BY SIMILARITY.
```

```
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    135    135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    140    140       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    159    159       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    163    163       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    188    188       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    189    189       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    199    199       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    209    209       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    246    246       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    253    253       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    288    288       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    307    307       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    366    366       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    372    372       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    402    402       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    408    408       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    412    412       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    418    418       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    423    423       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    460    460       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    475    475       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    622    622       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    627    627       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    636    636       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    648    648       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   867 AA;  98399 MW;  5F2310146B8E8680 CRC64;

  Query Match             100.0%;  Score 46;  DB 1;  Length 867;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db        264 KPVVSTQLLL 273


RESULT 7
ENV_HV1Z3
ID   ENV_HV1Z3      STANDARD;      PRT;   460 AA.
AC   P12491;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (Zaire 3 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11680;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86259728; PubMed=3014529;
RA   Willey R.W., Rutledge R.A., Dias S., Folks T., Theodore T.,
RA   Buckler C.E., Martin M.A.;
RT   "Identification of conserved and divergent domains within the
RT   envelope gene of the acquired immunodeficiency syndrome retrovirus.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:5038-5042(1986).
CC   -!- MISCELLANEOUS: THOUGH THIS SEQUENCE CONTAINS A COMPLETE ENV CODING
CC       REGION, INSERTION OF AN EXTRA NUCLEOTIDE CREATES A STOP CODON
CC       PRIOR TO THE NORMAL TERMINATION; THE AUTHORS SUGGEST THAT THIS
CC       VARIATION CAN ACCOUNT FOR THE LACK OF INFECTIVITY OF THIS CLONE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K03347; AAA45372.1; -.
DR   EMBL; K03347; AAA45373.1; -.
DR   HIV; K03347; ENV$Z3.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL       1     29
FT   CHAIN       30    460       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   DISULFID     53     73       BY SIMILARITY.
FT   DISULFID    118    202       BY SIMILARITY.
FT   DISULFID    125    193       BY SIMILARITY.
FT   DISULFID    130    149       BY SIMILARITY.
FT   DISULFID    215    244       BY SIMILARITY.
FT   DISULFID    225    236       BY SIMILARITY.
FT   DISULFID    293    326       BY SIMILARITY.
FT   DISULFID    372    439       BY SIMILARITY.
FT   DISULFID    379    412       BY SIMILARITY.
FT   DISULFID    397    404       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    135    135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    139    139       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    148    148       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    177    177       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    180    180       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    181    181       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    194    194       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    231    231       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    238    238       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    259    259       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    273    273       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    286    286       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    327    327       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    356    356       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    380    380       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    390    390       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    407    407       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    442    442       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   460 AA;  51297 MW;  27E97EB75C7EFF50 CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 460;
  Best Local Similarity   90.0%;  Pred. No. 0.24;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        249 RPVVSTQLLL 258


RESULT 8
ENV_HV1Y2
ID   ENV_HV1Y2      STANDARD;      PRT;   843 AA.
AC   P35961;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=36377;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93021387; PubMed=1404605;
RA   Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA   Shaw G.M.;
```

```
RT   "Complete nucleotide sequence, genome organization, and biological
RT   properties of human immunodeficiency virus type 1 in vivo: evidence
RT   for limited defectiveness and complementation.";
RL   J. Virol. 66:6587-6600(1992).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M93258; -; NOT_ANNOTATED_CDS.
DR   PIR; H44001; H44001.
DR   PDB; 1G9N; 27-DEC-00.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW   3D-structure.
FT   SIGNAL       1    29
FT   CHAIN       30   489       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN      490   843       TRANSMEMBRANE GLYCOPROTEIN.
FT   TRANSMEM   738   755       POTENTIAL.
FT   DISULFID    53    73       BY SIMILARITY.
FT   DISULFID   118   201       BY SIMILARITY.
FT   DISULFID   125   192       BY SIMILARITY.
FT   DISULFID   130   155       BY SIMILARITY.
FT   DISULFID   214   243       BY SIMILARITY.
FT   DISULFID   224   235       BY SIMILARITY.
FT   DISULFID   292   326       BY SIMILARITY.
FT   DISULFID   373   432       BY SIMILARITY.
FT   DISULFID   380   405       BY SIMILARITY.
FT   CARBOHYD    87    87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   129   129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   135   135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   138   138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   154   154       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   158   158       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   184   184       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   193   193       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   230   230       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   237   237       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   258   258       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   272   272       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   285   285       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   291   291       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   297   297       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   327   327       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   351   351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   381   381       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   389   389       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   395   395       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   400   400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   435   435       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   450   450       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   598   598       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   603   603       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   612   612       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   624   624       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   803   803       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   843 AA;  95648 MW;  C69DFD971C918B71 CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 843;
  Best Local Similarity   90.0%;  Pred. No. 0.43;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :||||||||||
Db        248 RPVVSTQLLL 257
```

```
RESULT 9
ENV_HV1ND
ID   ENV_HV1ND      STANDARD;      PRT;   846 AA.
AC   P18799;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11695;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90034200; PubMed=2806917;
RA   Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
RA   Hampe A., Chermann J.C.;
RT   "Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the
RT   human immunodeficiency virus.";
RL   Gene 81:275-284(1989).
CC   -!- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC       AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M27323; AAA44873.1; -.
DR   PIR; JQ0066; VCLJND.
DR   HIV; M27323; ENV$NDK.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL       1    29
FT   CHAIN       30   501       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN      502   846       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID    53    73       BY SIMILARITY.
FT   DISULFID   118   200       BY SIMILARITY.
FT   DISULFID   125   191       BY SIMILARITY.
FT   DISULFID   130   152       BY SIMILARITY.
FT   DISULFID   213   242       BY SIMILARITY.
FT   DISULFID   223   234       BY SIMILARITY.
FT   DISULFID   291   328       BY SIMILARITY.
FT   DISULFID   374   435       BY SIMILARITY.
FT   DISULFID   381   408       BY SIMILARITY.
FT   CARBOHYD    87    87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   129   129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   151   151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   179   179       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   182   182       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   229   229       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   236   236       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   257   257       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   271   271       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   284   284       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   290   290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   351   351       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   382   382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   388   388       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   392   392       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   395   395       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   401   401       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   438   438       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   451   451       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    452    452       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    601    601       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    606    606       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    615    615       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    627    627       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   846 AA;  96476 MW;  8A3B9DA527DE2E83 CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 846;
  Best Local Similarity   90.0%;  Pred. No. 0.43;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        247 RPVVSTQLLL 256


RESULT 10
ENV_HV1S1
ID   ENV_HV1S1      STANDARD;      PRT;   847 AA.
AC   P19550;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11691;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90347835; PubMed=2384920;
RA   Cheng-Mayer C., Quiroga M., Tung J.W., Dina D., Levy J.;
RT   "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT   macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL   J. Virol. 64:4390-4398(1990).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M65024; AAA45072.1; -.
DR   PDB; 1OBE; 15-MAY-97.
DR   HIV; M38428; ENV$SF162.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW   3D-structure.
FT   SIGNAL        1     29
FT   CHAIN        30    502       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN       503    847       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID     53     73       BY SIMILARITY.
FT   DISULFID    118    203       BY SIMILARITY.
FT   DISULFID    125    194       BY SIMILARITY.
FT   DISULFID    130    155       BY SIMILARITY.
FT   DISULFID    216    245       BY SIMILARITY.
FT   DISULFID    226    237       BY SIMILARITY.
FT   DISULFID    294    328       BY SIMILARITY.
FT   DISULFID    374    435       BY SIMILARITY.
FT   DISULFID    381    408       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    135    135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    154    154       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    186    186       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    232    232       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    260    260       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    299    299       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    329    329       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    382    382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    388    388       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    392    392       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    398    398       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    401    401       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    438    438       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    454    454       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    602    602       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    607    607       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    616    616       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    628    628       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   847 AA;  96135 MW;  0A901317FD7FF2AB CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 847;
  Best Local Similarity   90.0%;  Pred. No. 0.43;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        250 RPVVSTQLLL 259


RESULT 11
ENV_HV1W2
ID   ENV_HV1W2      STANDARD;      PRT;   847 AA.
AC   P05880;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11705;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86235450; PubMed=3012778;
RA   Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA   Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT   "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
RT   at risk for AIDS.";
RL   Science 232:1548-1553(1986).
CC   -!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC       BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC       WAS PERINATALLY INFECTED BY HER MOTHER.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M12507; AAB12990.1; -.
DR   HIV; M12507; ENV$WMJ2.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL        1     29
FT   CHAIN        30    501       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN       502    847       TRANSMEMBRANE GLYCOPROTEIN.
```

```
FT   DISULFID     53     73       BY SIMILARITY.
FT   DISULFID    118    202       BY SIMILARITY.
FT   DISULFID    125    193       BY SIMILARITY.
FT   DISULFID    130    152       BY SIMILARITY.
FT   DISULFID    215    244       BY SIMILARITY.
FT   DISULFID    225    236       BY SIMILARITY.
FT   DISULFID    293    326       BY SIMILARITY.
FT   DISULFID    372    435       BY SIMILARITY.
FT   DISULFID    379    408       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    134    134       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    140    140       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    151    151       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    155    155       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    183    183       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    184    184       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    194    194       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    231    231       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    238    238       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    259    259       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    273    273       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    286    286       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    292    292       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    327    327       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    334    334       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    350    350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    356    356       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    380    380       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    386    386       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    390    390       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    438    438       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    450    450       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    602    602       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    607    607       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    616    616       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    628    628       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   847 AA;  96466 MW;  CD1E33D73AA5BCAE CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 847;
  Best Local Similarity   90.0%;  Pred. No. 0.43;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        249 RPVVSTQLLL 258


RESULT 12
ENV_HV1JR
ID   ENV_HV1JR      STANDARD;      PRT;   848 AA.
AC   P20871;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11688;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Koyanagi S., Chen I.S.Y.;
RL   Submitted (DEC-1988) to the HIV data bank.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38429; AAB03749.1; -.
DR   PDB; 1CE4; 18-MAR-99.
DR   HIV; M38429; ENV$JRCSF.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW   3D-structure.
FT   SIGNAL       1     32
FT   CHAIN       33    503       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN      504    848       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID     53     73       BY SIMILARITY.
FT   DISULFID    118    203       BY SIMILARITY.
FT   DISULFID    125    194       BY SIMILARITY.
FT   DISULFID    130    154       BY SIMILARITY.
FT   DISULFID    216    245       BY SIMILARITY.
FT   DISULFID    226    237       BY SIMILARITY.
FT   DISULFID    294    328       BY SIMILARITY.
FT   DISULFID    374    437       BY SIMILARITY.
FT   DISULFID    381    410       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    134    134       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    137    137       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    153    153       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    157    157       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    185    185       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    195    195       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    228    228       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    239    239       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    260    260       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    274    274       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    287    287       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    293    293       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    299    299       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    329    329       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    336    336       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    352    352       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    382    382       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    388    388       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    392    392       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    403    403       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    440    440       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    453    453       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    603    603       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    608    608       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    617    617       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    629    629       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    808    808       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   848 AA;  96475 MW;  20767F51227EC3F3 CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 848;
  Best Local Similarity   90.0%;  Pred. No. 0.43;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        250 RPVVSTQLLL 259


RESULT 13
ENV_HV1B8
ID   ENV_HV1B8      STANDARD;      PRT;   851 AA.
AC   P04582;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
```

```
OX   NCBI_TaxID=11684;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85111123; PubMed=2578615;
RA   Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA   Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA   Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA   Lautenberger J.A., Papas T.S., Ghrayeb J., Chang N.T., Gallo R.C.,
RA   Wong-Staal F.;
RT   "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL   Nature 313:277-284(1985).
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02011; AAA44661.1; -.
DR   PDB; 1DDH; 13-JAN-99.
DR   PDB; 1HHG; 31-OCT-93.
DR   PDB; 1QO3; 02-JAN-00.
DR   PDB; 1SZT; 24-DEC-97.
DR   HIV; K02011; ENV$BH8.
DR   GlycoSuiteDB; P04582; -.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW   3D-structure.
FT   SIGNAL       1    30
FT   CHAIN       31   506       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN      507   851       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID    54    74       BY SIMILARITY.
FT   DISULFID   119   205       BY SIMILARITY.
FT   DISULFID   126   196       BY SIMILARITY.
FT   DISULFID   131   157       BY SIMILARITY.
FT   DISULFID   218   247       BY SIMILARITY.
FT   DISULFID   228   239       BY SIMILARITY.
FT   DISULFID   296   331       BY SIMILARITY.
FT   DISULFID   378   440       BY SIMILARITY.
FT   DISULFID   385   413       BY SIMILARITY.
FT   CARBOHYD    88    88       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   136   136       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   141   141       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   156   156       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   160   160       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   186   186       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   197   197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   230   230       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   234   234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   241   241       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   262   262       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   276   276       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   295   295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   301   301       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   332   332       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   339   339       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   356   356       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   386   386       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   392   392       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   401   401       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   443   443       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   458   458       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   606   606       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   611   611       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   620   620       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   632   632       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   669   669       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   745   745       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   811   811       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   851 AA;  96644 MW;  D16A3C90857785F1 CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 851;
  Best Local Similarity   90.0%;  Pred. No. 0.44;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        252 RPVVSTQLLL 261


RESULT 14
ENV_HV1BN
ID   ENV_HV1BN      STANDARD;      PRT;   852 AA.
AC   P12488;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11693;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89085613; PubMed=2789516;
RA   Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P.,
RA   Dandekar S.;
RT   "Biological and molecular characterization of human immunodeficiency
RT   virus (HIV-1BR) from the brain of a patient with progressive
RT   dementia.";
RL   Virology 168:79-89(1989).
CC   -!- MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS
CC       HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21098; AAA44221.1; -.
DR   PIR; A31667; VCLJBR.
DR   PDB; 1IM7; 23-OCT-02.
DR   HIV; M21098; ENV$BRVA.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal;
KW   3D-structure.
FT   SIGNAL       1    30
FT   CHAIN       31   507       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN      508   852       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID    54    74       BY SIMILARITY.
FT   DISULFID   119   205       BY SIMILARITY.
FT   DISULFID   126   196       BY SIMILARITY.
FT   DISULFID   131   155       BY SIMILARITY.
FT   DISULFID   218   247       BY SIMILARITY.
FT   DISULFID   228   239       BY SIMILARITY.
FT   DISULFID   296   330       BY SIMILARITY.
FT   DISULFID   376   439       BY SIMILARITY.
FT   DISULFID   383   412       BY SIMILARITY.
FT   CARBOHYD    49    49       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    88    88       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   135   135       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   138   138       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   154   154       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   158   158       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   CARBOHYD    197    197       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    234    234       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    241    241       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    262    262       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    276    276       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    289    289       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    295    295       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    301    301       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    354    354       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    360    360       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    384    384       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    390    390       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    396    396       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    442    442       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    456    456       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    607    607       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    612    612       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    621    621       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    633    633       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    670    670       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    812    812       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   852 AA;  97203 MW;  2BB866345DEC915F CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 852;
  Best Local Similarity   90.0%;  Pred. No. 0.44;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        252 RPVVSTQLLL 261


RESULT 15
ENV_HV1EL
ID   ENV_HV1EL      STANDARD;      PRT;   853 AA.
AC   P04581;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Envelope polyprotein GP160 precursor [Contains: Exterior membrane
DE   glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN   ENV.
OS   Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86245056; PubMed=2424612;
RA   Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
RT   "Genetic variability of the AIDS virus: nucleotide sequence analysis
RT   of two isolates from African patients.";
RL   Cell 46:63-74(1986).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; K03454; AAA44329.1; -.
DR   EMBL; A07108; CAA00616.1; -.
DR   HIV; K03454; ENV$ELI.
DR   InterPro; IPR000328; Env_GP41.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
DR   Pfam; PF00517; GP41; 1.
KW   AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW   Signal.
FT   SIGNAL        1     31       BY SIMILARITY.
FT   CHAIN        32    508       EXTERIOR MEMBRANE GLYCOPROTEIN.
FT   CHAIN       509    853       TRANSMEMBRANE GLYCOPROTEIN.
FT   DISULFID     53     73       BY SIMILARITY.
FT   DISULFID    118    206       BY SIMILARITY.
FT   DISULFID    125    197       BY SIMILARITY.
FT   DISULFID    130    154       BY SIMILARITY.
FT   DISULFID    219    248       BY SIMILARITY.
FT   DISULFID    229    240       BY SIMILARITY.
FT   DISULFID    297    330       BY SIMILARITY.
FT   DISULFID    376    442       BY SIMILARITY.
FT   DISULFID    383    416       BY SIMILARITY.
FT   CARBOHYD     87     87       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    129    129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    137    137       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    143    143       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    153    153       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    157    157       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    183    183       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    188    188       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    198    198       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    235    235       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    242    242       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    277    277       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    290    290       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    331    331       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    353    353       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    384    384       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    390    390       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    394    394       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    400    400       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    405    405       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    406    406       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    411    411       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    458    458       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    459    459       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    462    462       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    608    608       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    613    613       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    622    622       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    634    634       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   853 AA;  96721 MW;  F9CD864DAA0D07A5 CRC64;

  Query Match             93.5%;  Score 43;  DB 1;  Length 853;
  Best Local Similarity   90.0%;  Pred. No. 0.44;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              :|||||||||
Db        253 RPVVSTQLLL 262


Search completed: March 12, 2004, 14:24:18
Job time : 12 secs
```

OM protein - protein search, using sw model

Run on: March 12, 2004, 14:22:50 ; Search time 40 Seconds
(without alignments)
78.880 Million cell updates/sec

Title: PARKIN524.PEP
Perfect score: 46
Sequence: 1 kpvvstqlll 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	46	100.0	33	15	Q9QKI9	Q9qki9 human immun
2	46	100.0	35	15	Q9J4I4	Q9j4i4 human immun
3	46	100.0	74	15	Q9J4H8	Q9j4h8 human immun
4	46	100.0	85	15	Q8J9C9	Q8j9c9 human immun
5	46	100.0	85	15	Q8J9C8	Q8j9c8 human immun
6	46	100.0	95	15	Q90U89	Q90u89 human immun
7	46	100.0	99	15	Q9QT67	Q9qt67 human immun
8	46	100.0	99	15	Q9QT80	Q9qt80 human immun
9	46	100.0	99	15	O91506	O91506 human immun
10	46	100.0	99	15	Q79317	Q79317 human immun
11	46	100.0	99	15	Q9QT79	Q9qt79 human immun
12	46	100.0	99	15	O91507	O91507 human immun
13	46	100.0	99	15	O91505	O91505 human immun
14	46	100.0	99	15	O91502	O91502 human immun
15	46	100.0	99	15	Q9QT78	Q9qt78 human immun
16	46	100.0	99	15	O89210	O89210 human immun
17	46	100.0	99	15	Q9QT69	Q9qt69 human immun
18	46	100.0	99	15	Q9QT71	Q9qt71 human immun
19	46	100.0	99	15	O89206	O89206 human immun
20	46	100.0	99	15	Q9QT73	Q9qt73 human immun
21	46	100.0	99	15	Q9QT70	Q9qt70 human immun
22	46	100.0	99	15	O91500	O91500 human immun
23	46	100.0	99	15	Q9QT72	Q9qt72 human immun
24	46	100.0	99	15	O91501	O91501 human immun
25	46	100.0	99	15	Q79318	Q79318 human immun
26	46	100.0	99	15	Q9QT75	Q9qt75 human immun
27	46	100.0	99	15	O91504	O91504 human immun
28	46	100.0	99	15	O91503	O91503 human immun
29	46	100.0	99	15	Q9QT74	Q9qt74 human immun
30	46	100.0	99	15	Q9QT68	Q9qt68 human immun
31	46	100.0	99	15	Q9QT77	Q9qt77 human immun
32	46	100.0	99	15	O89196	O89196 human immun
33	46	100.0	100	15	Q9YN56	Q9yn56 human immun
34	46	100.0	101	15	O40202	O40202 human immun
35	46	100.0	106	15	Q8JAS7	Q8jas7 human immun
36	46	100.0	106	15	O73212	O73212 human immun
37	46	100.0	107	15	Q9YTT1	Q9ytt1 human immun
38	46	100.0	107	15	Q9WLD5	Q9wld5 human immun
39	46	100.0	107	15	O90590	O90590 human immun
40	46	100.0	107	15	Q9YTS6	Q9yts6 human immun
41	46	100.0	107	15	Q9YTS2	Q9yts2 human immun
42	46	100.0	107	15	Q9WLD3	Q9wld3 human immun
43	46	100.0	107	15	Q9YTS3	Q9yts3 human immun
44	46	100.0	107	15	Q9YTS4	Q9yts4 human immun
45	46	100.0	107	15	Q9YTS5	Q9yts5 human immun

ALIGNMENTS

RESULT 1
Q9QKI9
ID Q9QKI9 PRELIMINARY; PRT; 33 AA.
AC Q9QKI9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22202v3.11;
RX MEDLINE=99094949; PubMed=9878014;
RA Van Dyke R.B., Korber B.T., Popek E., Macken C., Widmayer S.M.,
RA Bardeguez A., Hansen I.C., Wiznia A., Luzuriaga K., Viscarello R.R.,
RA Wolinsky S., the Ariel Core Investigators;
RT "The Ariel Project: A prospective cohort study of maternal-child
RT transmission of human immunodeficiency virus type 1 in the era of
RT maternal antiretroviral therapy.";
RL J. Infect. Dis. 179:319-328(1999).
DR EMBL; AF112549; AAF13327.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW AIDS; Coat protein; Glycoprotein.
FT NON_TER 1 1
SQ SEQUENCE 33 AA; 3615 MW; 7E5F0E44BA535391 CRC64;

Query Match 100.0%; Score 46; DB 15; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KPVVSTQLLL 10
||||||||||

```
Db          13 KPVVSTQLLL 22


RESULT 2
Q9J4I4
ID   Q9J4I4       PRELIMINARY;      PRT;    35 AA.
AC   Q9J4I4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Truncated envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TB24;
RX   MEDLINE=20491829; PubMed=11035611;
RA   Collins K.R., Mayanja-Kizza H., Sullivan B.A., Quinones-Mateu M.E.,
RA   Toossi Z., Arts E.J.;
RT   "Greater diversity of HIV-1 quasispecies in HIV-infected individuals
RT   with active tuberculosis.";
RL   J. Acquir. Immune Defic. Syndr. 24:408-417(2000).
DR   EMBL; AF201787; AAF71854.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Envelope protein; Glycoprotein.
FT   NON_TER       1      1
SQ   SEQUENCE   35 AA;  3497 MW;  F424000992318A39 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db         18 KPVVSTQLLL 27


RESULT 3
Q9J4H8
ID   Q9J4H8       PRELIMINARY;      PRT;    74 AA.
AC   Q9J4H8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Truncated envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TB24;
RX   MEDLINE=20491829; PubMed=11035611;
RA   Collins K.R., Mayanja-Kizza H., Sullivan B.A., Quinones-Mateu M.E.,
RA   Toossi Z., Arts E.J.;
RT   "Greater diversity of HIV-1 quasispecies in HIV-infected individuals
RT   with active tuberculosis.";
RL   J. Acquir. Immune Defic. Syndr. 24:408-417(2000).
DR   EMBL; AF201793; AAF71860.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Envelope protein; Glycoprotein.
FT   NON_TER       1      1
SQ   SEQUENCE   74 AA;  8011 MW;  59A00D631924528A CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 0.037;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db         18 KPVVSTQLLL 27


RESULT 4
Q8J9C9
ID   Q8J9C9       PRELIMINARY;      PRT;    85 AA.
AC   Q8J9C9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ARBL69;
RA   Ceballos A., Rabinovich R.D., Avila M.M., Martinez Peralta L.,
RA   Gomez Carrillo M.;
RT   "Molecular study of an HIV-1 transmission chain 7 years after
RT   suspected events.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF490622; AAM90890.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      85     85
SQ   SEQUENCE   85 AA;  9372 MW;  70F92F1CE393D580 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 0.043;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          7 KPVVSTQLLL 16


RESULT 5
Q8J9C8
ID   Q8J9C8       PRELIMINARY;      PRT;    85 AA.
AC   Q8J9C8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ARBL68;
RA   Ceballos A., Rabinovich R.D., Avila M.M., Martinez Peralta L.,
RA   Gomez Carrillo M.;
RT   "Molecular study of an HIV-1 transmission chain 7 years after
RT   suspected events.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF490623; AAM90891.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
```

```
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      85     85
SQ   SEQUENCE   85 AA;  9232 MW;  6034854CF928A3F5 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 0.043;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          7 KPVVSTQLLL 16


RESULT 6
Q90U89
ID   Q90U89      PRELIMINARY;      PRT;    95 AA.
AC   Q90U89;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21192328; PubMed=11294662;
RA   Gunthard H.F., Havlir D.V., Fiscus S., Zhang Z.-Q., Eron J.,
RA   Mellors J., Gulick R., Frost S.D., Leigh Brown A.J., Schleif W.,
RA   Valentine F., Jonas L., Meibohm A., Ignacio C.C., Isaacs R.,
RA   Gamagami R., Emini E., Haase A., Richman D.D., Wong J.K.;
RT   "Residual human immunodeficiency virus (HIV) type 1 RNA and DNA in
RT   lymph nodes and HIV RNA in genital secretions and in cerebrospinal
RT   fluid after suppression of viremia for 2 years.";
RL   J. Infect. Dis. 183:1318-1327(2001).
DR   EMBL; AF337312; AAK56233.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      95     95
SQ   SEQUENCE   95 AA;  10507 MW;  D5A14913D1093609 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 0.048;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          5 KPVVSTQLLL 14


RESULT 7
Q9QT67
ID   Q9QT67      PRELIMINARY;      PRT;    99 AA.
AC   Q9QT67;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein gp120 C2V3 region (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=subtype E;
RX   MEDLINE=99408506; PubMed=10480629;
RA   Kato K., Shiino T., Kusagawa S., Sato H., Nohtomi K., Shibamura K.,
RA   Hien N.T., Chi P.K., Lien T.X., Anh M.H., Long H.T.,
RA   Bunyaraksyotin G., Fukushima Y., Honda M., Wasi C., Yamazaki S.,
RA   Nagai Y., Takebe Y.;
RT   "Genetic similarity of HIV Type 1 subtype E in a recent outbreak among
RT   injecting drug users in Northern Vietnam to strains in Guangxi
RT   province of Southern China.";
RL   AIDS Res. Hum. Retroviruses 15:1157-1168(1999).
DR   EMBL; AB025097; BAA83669.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      99     99
SQ   SEQUENCE   99 AA;  11109 MW;  AAB1BA8E0DB38E13 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12


RESULT 8
Q9QT80
ID   Q9QT80      PRELIMINARY;      PRT;    99 AA.
AC   Q9QT80;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein gp120 C2V3 region (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=subtype E;
RX   MEDLINE=99408506; PubMed=10480629;
RA   Kato K., Shiino T., Kusagawa S., Sato H., Nohtomi K., Shibamura K.,
RA   Hien N.T., Chi P.K., Lien T.X., Anh M.H., Long H.T.,
RA   Bunyaraksyotin G., Fukushima Y., Honda M., Wasi C., Yamazaki S.,
RA   Nagai Y., Takebe Y.;
RT   "Genetic similarity of HIV Type 1 subtype E in a recent outbreak among
RT   injecting drug users in Northern Vietnam to strains in Guangxi
RT   province of Southern China.";
RL   AIDS Res. Hum. Retroviruses 15:1157-1168(1999).
DR   EMBL; AB025084; BAA83656.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      99     99
SQ   SEQUENCE   99 AA;  10954 MW;  504F1205EBF2D1C5 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12
```

```
RESULT 9
O91506
ID   O91506      PRELIMINARY;      PRT;    99 AA.
AC   O91506;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=subtype E;
RX   MEDLINE=99146664; PubMed=10024059;
RA   Kusagawa S., Sato H., Kato K., Nohtomi K., Shiino T., Samrith C.,
RA   Leng H.B., Phalla T., Heng M.B., Takebe Y.;
RT   "HIV type 1 env subtype E in Cambodia.";
RL   AIDS Res. Hum. Retroviruses 15:91-94(1999).
DR   EMBL; AB013126; BAA33687.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      99     99
SQ   SEQUENCE   99 AA;  10927 MW;  773564329BF2D1DB CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12


RESULT 10
Q79317
ID   Q79317      PRELIMINARY;      PRT;    99 AA.
AC   Q79317;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JNIH2T;
RA   Takebe Y., Fujiwara Y., Sato H., Oka S., Pau C.P., Imai M., Ou C.Y.,
RA   Weniger B.G., Yamazaki S.;
RT   "Identification of Thai HIV-1 Variants in Japan (provisional title).";
RL   Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JNIH2T;
RA   Weniger B.G., Takebe Y., Ou C.Y., Yamazaki S.;
RT   "The molecular epidemiology of HIV in Asia.";
RL   AIDS 0:0-0(1994).
DR   EMBL; L32085; AAC37879.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
SQ   SEQUENCE   99 AA;  11012 MW;  A07B123AE109DF23 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12


RESULT 11
Q9QT79
ID   Q9QT79      PRELIMINARY;      PRT;    99 AA.
AC   Q9QT79;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein gp120 C2V3 region (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=subtype E;
RX   MEDLINE=99408506; PubMed=10480629;
RA   Kato K., Shiino T., Kusagawa S., Sato H., Nohtomi K., Shibamura K.,
RA   Hien N.T., Chi P.K., Lien T.X., Anh M.H., Long H.T.,
RA   Bunyaraksyotin G., Fukushima Y., Honda M., Wasi C., Yamazaki S.,
RA   Nagai Y., Takebe Y.;
RT   "Genetic similarity of HIV Type 1 subtype E in a recent outbreak among
RT   injecting drug users in Northern Vietnam to strains in Guangxi
RT   province of Southern China.";
RL   AIDS Res. Hum. Retroviruses 15:1157-1168(1999).
DR   EMBL; AB025085; BAA83657.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      99     99
SQ   SEQUENCE   99 AA;  10978 MW;  B544B34E46A5A2C3 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12


RESULT 12
O91507
ID   O91507      PRELIMINARY;      PRT;    99 AA.
AC   O91507;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=subtype E;
RX   MEDLINE=99146664; PubMed=10024059;
```

```
RA   Kusagawa S., Sato H., Kato K., Nohtomi K., Shiino T., Samrith C.,
RA   Leng H.B., Phalla T., Heng M.B., Takebe Y.;
RT   "HIV type 1 env subtype E in Cambodia.";
RL   AIDS Res. Hum. Retroviruses 15:91-94(1999).
DR   EMBL; AB013127; BAA33688.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      99     99
SQ   SEQUENCE   99 AA;  10966 MW;  92C67C25FE0731C0 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12


RESULT 13
O91505
ID   O91505      PRELIMINARY;      PRT;    99 AA.
AC   O91505;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=subtype E;
RX   MEDLINE=99146664; PubMed=10024059;
RA   Kusagawa S., Sato H., Kato K., Nohtomi K., Shiino T., Samrith C.,
RA   Leng H.B., Phalla T., Heng M.B., Takebe Y.;
RT   "HIV type 1 env subtype E in Cambodia.";
RL   AIDS Res. Hum. Retroviruses 15:91-94(1999).
DR   EMBL; AB013125; BAA33686.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      99     99
SQ   SEQUENCE   99 AA;  11007 MW;  3BD7758D878C2357 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12


RESULT 14
O91502
ID   O91502      PRELIMINARY;      PRT;    99 AA.
AC   O91502;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=subtype E;
RX   MEDLINE=99146664; PubMed=10024059;
RA   Kusagawa S., Sato H., Kato K., Nohtomi K., Shiino T., Samrith C.,
RA   Leng H.B., Phalla T., Heng M.B., Takebe Y.;
RT   "HIV type 1 env subtype E in Cambodia.";
RL   AIDS Res. Hum. Retroviruses 15:91-94(1999).
DR   EMBL; AB013121; BAA33683.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      99     99
SQ   SEQUENCE   99 AA;  10998 MW;  02DBE55EDA17FBB5 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12


RESULT 15
Q9QT78
ID   Q9QT78      PRELIMINARY;      PRT;    99 AA.
AC   Q9QT78;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Envelope glycoprotein gp120 C2V3 region (Fragment).
GN   ENV.
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=subtype E;
RX   MEDLINE=99408506; PubMed=10480629;
RA   Kato K., Shiino T., Kusagawa S., Sato H., Nohtomi K., Shibamura K.,
RA   Hien N.T., Chi P.K., Lien T.X., Anh M.H., Long H.T.,
RA   Bunyaraksyotin G., Fukushima Y., Honda M., Wasi C., Yamazaki S.,
RA   Nagai Y., Takebe Y.;
RT   "Genetic similarity of HIV Type 1 subtype E in a recent outbreak among
RT   injecting drug users in Northern Vietnam to strains in Guangxi
RT   province of Southern China.";
RL   AIDS Res. Hum. Retroviruses 15:1157-1168(1999).
DR   EMBL; AB025086; BAA83658.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0019031; C:viral envelope; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000777; GP120.
DR   Pfam; PF00516; GP120; 1.
KW   AIDS; Coat protein; Glycoprotein.
FT   NON_TER       1      1
FT   NON_TER      99     99
SQ   SEQUENCE   99 AA;  10969 MW;  5050E7576E0731C5 CRC64;

  Query Match             100.0%;  Score 46;  DB 15;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 0.05;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KPVVSTQLLL 10
              ||||||||||
Db          3 KPVVSTQLLL 12
```

Search completed: March 12, 2004, 14:25:11
Job time : 42 secs